Legal Medicine 83 (2026) 102848

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

The forensic evaluation of torture in asylum seekers: international literature comparison and inter-rater agreement analysis☆

Antonina Argo [a], Giuseppe Davide Albano [a,*], Giuseppe Pulin [b], Davide Ferorelli [b], Biagio Solarino [b], Mauro Midiri [a], Maria Puntarello [a], Luca Ferrero [a], Corinne La Spina [a], Valeria Tullio [a], Graziana Calascibetta [a], Serena Geraci [a], Ginevra Malta [a], Stefania Zerbo [a]

[a] Institute of Legal Medicine, Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Palermo, Italy
[b] Interdisciplinary Department of Medicine, Section of Legal Medicine, University of Bari, Piazza Giulio, Cesare 11, 70100 Bari, Italy

ARTICLE INFO

Keywords:
Torture
Asylum seekers
Forensic medical certification
Istanbul protocol
Inter-rater agreement
Literature comparison

ABSTRACT

Asylum seekers often report exposure to persecution in their home countries, during transit, or while in detention. The primary purpose of a forensic medical certification is the objective documentation of physical and psychological findings associated with torture, abuse, or other forms of inhumane treatment. The study had two main objectives: (1) to analyze the demographic, clinical, and medico-legal data of individuals referred to the Specialized Outpatient Service for Survivors of Torture (SVT) at the Policlinic Hospital of the University of Palermo and (2) to evaluate how professional experience and specialized training shape the formulation of compatibility judgments according to the Istanbul Protocol, through a multicenter inter-rater analysis. This is a retrospective observational study that analyzed all consecutive forensic medical evaluations of asylum seekers at the Institute of Legal Medicine, University of Palermo, Italy, between January 1, 2018, and December 31, 2024. In total, 324 patients were examined at the institute's outpatient clinic for survivors of torture. Our findings highlight the essential role of specialized forensic experience in the medico-legal assessment of asylum seekers reporting torture and ill-treatment. Comparative analysis across multiple international centers confirmed that physical trauma, particularly blunt injury, was the most commonly documented form of abuse. The study demonstrated high inter-rater agreement for injury identification, classification, and recognizing associated physical and psychological sequelae. However, discrepancies emerged in assigning consistency levels under the Istanbul Protocol, highlighting the need for structured training and standardized protocols to ensure the accuracy, reproducibility, and evidentiary value of forensic assessments.

1. Introduction

In recent years, migration to Europe has markedly increased [1]. According to the European Agency for Asylum, 1.14 million applications for international protection were lodged across European countries in 2023, an 18% increase compared to 2022. Syrians constituted the largest applicant group in 2023, with a 38% increase compared to 2022. Among the 27 European Union Member States plus Norway, Switzerland, Iceland, and Liechtenstein (EU+), Germany remained the leading destination for asylum seekers, receiving nearly one-third of all applications, followed by France and Italy. Italy, particularly Sicily, received the highest proportion of migrants arriving via the Mediterranean due to its geographic location [2]. Many applicants for international protection report having experienced persecution in their countries of origin. Forensic assessments are therefore essential to support these claims and to guide decisions on refugee status (political asylum) or, if the criteria are not fulfilled, on subsidiary protection or temporary permits. According to official national data from the Department of Civil Liberties and Immigration of the Ministry of the Interior between 2018 and 2022, the most recent published statistical data, a total of 262,240 applications

☆ This article is part of a special issue entitled: 'Proceedings Congress Forensic Science' published in Legal Medicine.

* Corresponding author at: Institute of Legal Medicine, Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, Via del Vespro 129, 90100 Palermo Italy.

E-mail addresses: antonella.argo@unipa.it (A. Argo), giuseppedavide.albano@unipa.it (G.D. Albano), davide.ferorelli@uniba.it (D. Ferorelli), biagio.solarino@uniba.it (B. Solarino), mauro.midiri@unipa.it (M. Midiri), maria.puntarello@community.unipa.it (M. Puntarello), corinne.laspina@unipa.it (C.L. Spina), valeria.tullio@unipa.it (V. Tullio), graziana.calascibetta@community.unipa.it (G. Calascibetta), ginevra.malta@unipa.it (G. Malta), stefania.zerbo@unipa.it (S. Zerbo).

https://doi.org/10.1016/j.legalmed.2026.102848
Received 7 August 2025; Received in revised form 28 September 2025; Accepted 28 March 2026
Available online 21 April 2026

were submitted by adults and 8,017 by unaccompanied minors [3]. A substantial proportion of immigrants in Italy come from regions afflicted by conflict or oppressive regimes. Many of them apply for asylum, reporting experiences of persecution in their countries, during their journey, or in detention centers where they are vulnerable to abuse. In response to such allegations, forensic assessments are often requested by local committees responsible for determining eligibility for international protection under the Geneva Convention. A forensic medical certification plays a critical role in the asylum process [4–7]. This certification, when conducted in compliance with international standards, can provide critical evidence of persecution, either in their home country or during their journey [8–10]. The primary aim of a forensic medical certification is the objective documentation of physical and psychological findings associated with persecution, torture, abuse, or other forms of inhumane treatment. Such forensic evidence is essential in substantiating the asylum seeker's allegation of having a "well-founded fear of persecution" as required under international law [9,11–13]. Furthermore, the certification may identify severe disabling conditions or chronic illnesses requiring ongoing and specialized medical care [14,15]. It also provides prognostic information, enabling assessment of the potential health consequences of repatriation for the applicants [16,17]. The methodology of forensic medical evaluation primarily consists of observation, identification, and description of observable signs and clinical manifestations [18]. Equally important is the formulation of differential diagnoses based on the findings to determine whether they can be attributed to natural, induced, traumatic, or non-accidental causes [19]. The forensic medical examiner is obligated to document and interpret evidence objectively to describe the observed injuries, with the aim of providing a forensic medical opinion that accurately reflects the reality [20,21]. The purpose of examining the injury is to determine its nature, characteristics, and the methods of physical torture alleged by the survivors. Furthermore, the evaluation seeks to document all relevant elements to reconstruct the manner and timing of the traumatic event. The Istanbul Protocol, a manual developed under the patronage of the United Nations (UN) for health care professionals assisting victims of torture, outlined the appropriate procedures for conducting physical examinations, documenting evidence of torture, and assigning degrees of consistency between the survivor's narrative and the observed type of lesions [22]. In this retrospective analysis of forensic assessments of asylum seekers, we emphasize that the work of our institute has been strengthened through a collaborative protocol established with Médecins Sans Frontières, which supports our fieldwork and ensures the involvement of an intercultural mediator, thereby enhancing communication and cultural understanding [23]. The objectives of our study are as follows:

- To analyze and evaluate the demographic, clinical, and medico-legal data of individuals referred to the Specialized Outpatient Service for Survivors of Torture (SVT) at the Policlinic Hospital of the University of Palermo. This retrospective review of forensic medical reports aims to identify prevalent injury patterns and types, reported methods of torture, and corresponding physical and psychological outcomes. Findings will be compared with data from similar international centers to highlight statistically significant differences that may reflect regional or population-based variations.
- To assess the relevance of professional experience and specific training on the formulation of compatibility judgments according to the Istanbul Protocol, through a multicenter inter-rater analysis.

2. Materials and methods

This retrospective observational study analyzed all forensic medical evaluations conducted for asylum seekers at the Institute of Legal Medicine, University of Palermo (Italy) between January 1, 2018, and December 31, 2024. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and approved by the Bioethics Committee of the University of Palermo (Approval No. 137/2023).

2.1. Forensic assessment framework

All forensic assessments were conducted by physicians with specialized training in forensic medicine and documented expertise in the medico-legal evaluation of victims of violence and abuse, supported by a cultural-linguistic mediator and a forensic psychologist, ensuring accurate communication and trauma-informed care.

The methodology adhered to the principles and recommendations of the Istanbul Protocol (UN, 2022 revision) [22], the internationally recognized guideline for the documentation and investigation of torture and other cruel, inhuman, or degrading treatment or punishment. The protocol provided a structured framework for the clinical interview and physical examination, aiming to identify physical and psychological findings consistent with torture.

The methodology adhered to current medico-legal literature on vulnerability assessment among asylum seekers [11]. During the anamnesis phase, a specifically tailored medical chart was prefilled to systematically collect the following information: sociodemographic data, country of origin, reason for departure, migratory trajectory, detention history (including duration and location), reported torture methods, anatomical location and type of injuries, long-term physical, psychological, or functional sequelae, presence of permanent findings, and indicators of female genital mutilation (FGM).

2.2. Data collection and analysis

In 2025, all data were retrospectively reviewed and anonymized to minimize observer and selection bias. To contextualize our findings, a non-systematic comparative analysis was conducted using retrospective single-center studies from the international medico-legal literature. Targeted searches of the PubMed and Google Scholar databases using the terms "medico-legal evaluation," "forensic evaluation," "torture," and "asylum seekers" identified five relevant studies with homogeneous methodologies. These studies reported evaluations from Italy (Brescia) [24], France (Montpellier and Nantes) [25,26], and Sri Lanka [27].

Descriptive and inferential statistical analyses were conducted. Chi-square tests were applied to identify statistically significant differences in the prevalence of specific torture methods and outcomes, including blunt force trauma, sharp force trauma, burns, gunshot injuries, electrical injuries, and FGM.

2.3. Interobserver variability assessment

To assess the interobserver reliability and reproducibility of forensic evaluations, a structured interobserver analysis was conducted on a subsample of 103 cases from 2022 to 2023. This analysis included cases in which the observers had not participated in the examination, employing a methodology adapted from Franceschetti et al. [28]. The forensic medical records for this subset included full clinical histories and physical findings.

Three forensic physicians independently reviewed the documentation process and completed a standardized forensic report:

OF1: A resident in forensic medicine, trained in a center without specialized training in torture assessment.
OF2: A resident with prior experience ($\geq$50 evaluations), trained in a center with specialized training in torture documentation.
OF3: A senior forensic expert (>20 years of experience, >300 evaluations), designated as the reference evaluator.

Each evaluator assessed:

- Demographics and country of origin.

- Alleged torture methods and contexts, classified according to the Istanbul Protocol.
- Consistency between allegations and observed findings, rated on a five-point scale (not consistent, consistent with, highly consistent, typical of, diagnostic of torture).

To assess inter-rater agreement, Fleiss' K statistics were calculated for the five consistency categories.

2.4. Standardized data collection instrument

A comprehensive, mixed-format reporting tool was employed to enhance the reproducibility. It comprised:

- Anatomical injury mapping with binary indicators (yes/no) by body region (head/neck, trunk, limbs, and perineal region).
- Sequelae identification with binary responses (yes/no) for scars, functional impairments, sensory deficits, and psychological consequences.
- Torture method indicators with binary responses (yes/no) for trauma types (thermal, blunt, sharp, firearm, and electrical).
- Consistency grading: five-level classification according to Istanbul Protocol guidelines.

2.5. Informed consent and ethics approval

All forensic assessments were conducted in compliance with ethical and human rights principles. Written informed consent was obtained from each individual before examination, with a qualified cultural mediator present to ensure clear understanding and communication throughout the process. The assessments adhered to the Declaration of Helsinki (1975, revised 2013). The retrospective analysis of the collected data was conducted as part of a research project approved by the Bioethics Committee of the University of Palermo (Approval No. 137/2023, issued on February 23, 2023).

3. Results

3.1. Sample characteristics

Between January 1, 2018, and December 31, 2024, a total of 324 patients underwent physical examination for torture survivors at the outpatient clinic of the Institute of Legal Medicine of the University Hospital of Palermo. The cohort comprised 269 men (83%) and 55 women (17%), with a mean age of 26.63 years (SD ± 9.15).

Most of the assessed individuals originated from Sub-Saharan African countries, including Burkina Faso, Benin, Cameroon, Côte d'Ivoire, Chad, Guinea, Eritrea, Ethiopia, Guinea-Bissau, Guinea-Conakry, Gambia, Ghana, Liberia, Mali, Mauritius, Nigeria, Senegal, Sierra Leone, Somalia, Sudan, and South Sudan, accounting for 260 cases (80%). Migrants from Northern African countries (Egypt, Tunisia, and Libya) accounted for 23 cases (7%), while individuals from Bangladesh accounted for 38 cases (12%). Pakistan and Sri Lanka accounted for 3 cases (1%).

Consistency between allegations and findings was rated as consistent (n = 253, 78%), highly consistent (n = 30, 9%), typical (n = 30, 9%), specific (n = 8, 3%), and not consistent (n = 3, 1%). Injuries were most frequently located on the lower limbs (n = 227, 70%) and upper limbs (n = 217, 67%). Trunk injuries were observed in 178 patients (55%), and head and neck injuries in 152 (47%). Perineal injuries were documented in 32 cases (10%). Permanent physical and psychological findings were highly prevalent. Scars were observed in 305 individuals (94%), musculoskeletal injuries in 84 (26%), and neuropsychiatric conditions were reported in 321 cases (99%). Sensory impairments and genitourinary disorders were noted in 39 individuals (12%). A history of detention was reported by 230 individuals (71%), predominantly in Libya. Regarding the interval between the reported acts of torture and the subsequent forensic evaluation, the distribution was as follows: 40 individuals (12%) were evaluated within < 6 months; 45 (14%) within 6–12 months; 62 (19%) within 12–24 months; 43 (13%) within 24–36 months; 26 (8%) within 36–48 months; 19 (6%) within 48–60 months; and 47 (15%) after >60 months. Notably, 42 individuals (13%) could not specify the timing of the torture episode. Reported reasons for fleeing included ethnic persecution (n = 51, 16%), religious persecution (n = 24, 7%), political persecution (n = 63, 19%), economic hardship (n = 86, 27%), war or armed conflict (n = 39, 12%), domestic abuse (n = 85, 27%), discrimination based on sexual orientation (n = 9, 3%), and forced marriage practices (n = 23, 3%). Regarding injury types, injuries consistent with blunt force trauma, such as kicking, punching, whipping, or beatings with sticks or rods, were the most common (n = 288, 88.9%), followed by sharp injuries, typically caused by knives, awls, machetes, or bayonets (n = 116, 35.8%); thermal trauma, including cigarette burns, heated metal rods, acid burns, or so-called "fuel burns" (n = 88, 27.2%) [29,30]; firearm-related injuries, caused by pistols, rifles, or explosive devices (n = 28, 8.6%); injuries from electrical torture, such as electrified wires or tasers (n = 24 patients, 7.4%); and FGM (n = 36, 11.1%) (Table 1).

It should be noted that percentages related to reported reasons for fleeing, anatomical distribution of injuries, injury types, and permanent sequelae may exceed 100%, since these categories are not mutually exclusive. Individuals frequently reported many reasons for fleeing from their country. Patients also exhibited injuries at multiple sites and more than one type of long-term sequelae attributable to torture.

Table 1
This table presents key characteristics of the asylum seekers who underwent forensic medical assessment for alleged torture, including age, sex, country of origin, reasons for departure, history of detention, reported methods of torture, anatomical distribution of injuries, and long-term physical or psychological sequelae.

Variable	Category	N (%)
Total patients	—	324
Sex	Men	269 (83.0%)
	Women	55 (17.0%)
Mean age (SD)	—	26.63 (±9.15) years
Time since torture	<6 months	40 (12%)
	6–12 months	45 (14%)
	12–24 months	62 (19%)
	24–36 months	43 (13%)
	36–48 months	26 (8%)
	48–60 months	19 (6%)
	>60 months	47 (15%)
	Unknown	42 (13%)
Detention	Yes	230 (71%)
	No	94 (29%)
Reported reasons for fleeing	Ethnic persecution	51 (16%)
	Religious persecution	24 (7%)
	Political persecution	63 (19%)
	Economic hardship	86 (27%)
	War or armed conflict	39 (12%)
	Domestic abuse	85 (27%)
	Discrimination (sexual orientation)	9 (3%)
	Forced marriage	23 (7%)
Types of injury	Blunt force trauma	288 (88.9%)
	Sharp object injuries	116 (35.8%)
	Thermal trauma	88 (27.2%)
	Firearm-related injuries	28 (8.6%)
	Electrical torture	24 (7.4%)
	Female genital mutilation	36 (11.1%)
Permanent findings	Scars	305 (94%)
	Musculoskeletal injuries	84 (26%)
	Psychological and psychiatric sequelae	321 (99%)
	Other	39 (12%)

Furthermore, FGM was documented in 36 out of the total 324 patients (11.1%), but when considering only the female subgroup (n = 55), the prevalence of FGM rises to 65.5%.

3.2. Other experiences comparison

To contextualize our findings, a total of 1,636 medico-legal cases involving allegations of torture across five centers, Palermo (n = 324), Brescia (n = 185), Montpellier (n = 495), Nantes (n = 532), and a sample from Sri Lanka (n = 100), were reviewed.

Burn injuries were reported with variable prevalence. The highest incidence was observed in the Sri Lanka cohort (n = 57, 57%), followed by Palermo (n = 88, 27.2%), Montpellier (n = 100, 20.2%), Brescia (n = 35, 18.9%), and Nantes (n = 88, 16.5%). Blunt force trauma was the most prevalent type of injury in all centers, as in Sri Lanka (n = 95 cases, 95%), Palermo (n = 288, 88.9%), Brescia (n = 164, 88.6%), Nantes (n = 448 cases, 84.2%), and Montpellier (n = 414, 83.6%). Regarding injuries from sharp objects, the highest proportion was recorded in Brescia (n = 84 cases, 45.4%), followed by Palermo (n = 116 cases, 35.8%), Montpellier (n = 146 cases, 29.5%), Nantes (n = 158 cases, 29.7%), and Sri Lanka (n = 21 cases, 21%). Gunshot wounds were less frequently observed, and their prevalence was similar across the European centers: Palermo (n = 28, 8.6%), Brescia (n = 28 cases, 15.1%), Montpellier (n = 28, 5.7%), and Nantes (n = 31, 5.8%). However, they were not reported in the Sri Lankan sample. Electrical injuries were reported in the Sri Lanka center (n = 8 cases, 8%), Palermo (n = 24, 7.4%), Brescia (n = 20, 10.8%), Nantes (n = 29 cases, 5.5%), and Montpellier (n = 5 cases, 1%). FGM was reported in the Palermo center (n = 36, 11.1%) and Montpellier (n = 52 cases, 10.5%). Data on FGM were not reported in the remaining centers (Table 2).

To assess the distribution of injury types across different medico-legal centers, a chi-square goodness-of-fit test was applied. Table 3 presents expected frequencies for each injury type and center, calculated from the overall distribution of injuries in the total sample. These expected values correspond to the anticipated frequencies of injuries that would be anticipated, assuming the type of injury was distributed proportionally across all centers.

Table 3 also presents the standardized residuals derived from the chi-square test. These values quantify both the magnitude and direction of the deviation between observed and expected frequencies. Positive residuals indicate a higher-than-expected occurrence of a specific injury at a specific center, while negative residuals indicate a lower-than-expected frequency. Residuals with absolute values greater than 2 are typically considered statistically significant deviations.

Sri Lanka exhibited a markedly higher-than-expected number of burn injuries (+5.89) and a significantly lower-than-expected number of sharp force injuries (–2.75). Brescia exhibited higher-than-expected frequencies for injuries due to sharp objects (+1.75), gunshot wounds (+3.28), and electrical shock (+2.56), indicating a distinct injury profile compared to the other centers. Conversely, Montpellier reported a significantly lower-than-expected number of electrical injuries (–3.85).

Table 3
This table presents the expected frequencies and standardized residuals for each injury category by center, calculated from the overall sample distribution. Standardized residuals represent the extent to which observed values deviate from expected ones. Values ≥ |2| are typically regarded as statistically significant over- or under-representation (SR: standardized residual).

Center	Burns (E)/SR	Blunt Force (E)/SR	Sharp Object (E)/SR	Gunshot (E)/SR	Electrical (E)/SR
Palermo	80.0/+0.90	306.2/–1.04	114.1/+0.18	25.0/+0.60	18.7/+1.23
Brescia	48.7/–1.96	186.3/–1.64	69.4/+1.75	15.2/+3.28	11.4/+2.56
Montpellier	101.9/–0.19	390.1/+1.21	145.4/+0.05	31.8/–0.68	23.8/–3.85
Nantes	110.9/–2.17	424.4/+1.14	158.2/–0.01	34.6/–0.62	25.9/+0.61
Sri Lanka	26.6/+5.89	101.9/–0.68	38.0/–2.75	8.3/–	6.2/+0.71

3.3. Inter-observer variability

3.3.1. Agreement on the distribution of injuries

The degree of inter-rater agreement on the anatomical distribution of injuries was exceptionally strong. Cohen's kappa coefficient, calculated across all three evaluator pairs, was 1.000, indicating strong agreement. Similarly, Fleiss' kappa, used to assess overall agreement among all three observers, was also 1.000. In every case, all evaluators consistently identified the presence of findings in the same anatomical regions. This complete concordance may reflect the fact that all three physicians, including OF1, a trainee, had prior experience in forensic traumatology.

3.3.2. Agreement on the etiology of findings

Inter-rater agreement on the etiological classification of the injuries was strong across all trauma types, with some variability depending on the type of lesion. For thermal trauma, the mean kappa was 0.970, indicating near-perfect agreement. Pairwise values were: OF1 vs. OF2: 0.955; OF1 vs. OF3: 0.955; and OF2 vs. OF3: 1.000. For blunt force trauma, all evaluator pairs demonstrated perfect agreement (kappa = 1.000). For electrical trauma, the mean kappa was 0.882. Pairwise values were: OF1 vs. OF2: 0.904; OF1 vs. OF3: 0.823; and OF2 vs. OF3: 0.918. For sharp force trauma, the mean kappa was 0.885. Pairwise values were: OF1 vs. OF2: 0.869; OF1 vs. OF3: 0.828; and OF2 vs. OF3: 0.958. For firearm injuries, the mean kappa was 0.882. Pairwise values were: OF1 vs. OF2: 1.000; OF1 vs. OF3: 0.823; and OF2 vs. OF3: 0.823. These values reflect substantial to almost perfect agreement among evaluators in determining the likely mechanism of injury.

3.3.3. Agreement on the nature of sequelae

Evaluator agreement regarding the nature of the sequelae was also consistently high. For the presence of functional impairment, the mean kappa was 0.815. Pairwise values were: OF1 vs. OF2: 0.753; OF1 vs. OF3: 0.727; and OF2 vs. OF3: 0.966. For the presence of scarring, the mean kappa was 1.000, indicating perfect agreement. For the presence of sensory impairment, the mean kappa was 0.829. Pairwise values were: OF1 vs. OF2: 0.779; OF1 vs. OF3: 0.748; and OF2 vs. OF3: 0.960. For the presence of psychological sequelae, the mean kappa was 1.000,

Table 2
This table presents the absolute and relative frequencies of major injury categories (burns, blunt force, sharp object injuries, gunshots, electrical injuries, and female genital mutilation) documented during forensic medical evaluations conducted at five centers (FGM: female genital mutilations).

Center	Total Cases	Burns (n, %)	Blunt Force (n, %)	Sharp Object (n, %)	Gunshot (n, %)	Electrical Shock (n, %)	FGM (n, %)
Palermo	324	88 (27.2%)	288 (88.9%)	116 (35.8%)	28 (8.6%)	24 (7.4%)	36 (11.1%)
Brescia	185	35 (18.9%)	164 (88.6%)	84 (45.4%)	28 (15.1%)	20 (10.8%)	Not reported
Montpellier	495	100 (20.2%)	414 (83.6%)	146 (29.5%)	28 (5.7%)	5 (1%)	52 (10.5%)
Nantes	532	88 (16.5%)	448 (84.2%)	158 (29.7%)	31 (5.8%)	29 (5.5%)	Not reported
Sri Lanka	100	57 (57%)	95 (95%)	21 (21%)	Not reported	8 (8%)	Not reported

indicating complete agreement. Overall, these results indicate strong reliability in identifying both physical and psychological sequelae of torture.

3.3.4. Agreement on the degree of compatibility with alleged torture

To assess agreement on the degree of compatibility between clinical findings and the reported acts of torture—according to the Istanbul Protocol—a five-point ordinal scale was employed, ranging from:

1 = Not compatible.
2 = Compatible.
3 = Highly compatible.
4 = Typical.
5 = Specific.

Pairwise analysis of weighted kappa revealed a marked variability:

OF1 vs. OF2: 0.140 (poor agreement);
OF1 vs. OF3: 0.105 (poor agreement);
OF2 vs. OF3: 0.686 (moderate agreement).

These findings indicate that, although agreement was high on injury identification and classification, greater divergence than this agreement emerged in the interpretive evaluation of compatibility, particularly involving OF1. This reinforces the hypothesis that experience and specialized training in the forensic assessment of torture significantly enhance the consistency and reliability of evaluative judgments. The closer alignment between OF2 (a trained forensic physician in the field) and OF3 (a senior expert) indicates that evaluative consistency improves with specialization.

4. Discussion

The findings of this study align with the existing literature and data from other international centers conducting forensic assessments of torture among asylum seekers. The predominance of young adult men from Sub-Saharan Africa, the prevalence of blunt force trauma, and the high proportion of psychological and sexual abuse among female victims are frequent patterns reported across multiple cohorts, including Montpellier, Brescia, Nantes, and Sri Lanka.

Our findings indicate that blunt force trauma was the most prevalent type of injury (88.9%), aligning closely with findings from Montpellier (88.5%) and Sri Lanka (95%), highlighting this injury as a hallmark of torture. Burn injuries, indicative of deliberate infliction of pain, were also frequently reported in Palermo (27.2%), ranking second after Sri Lanka (57%). Sharp injuries were notably higher in Brescia (45.4%) than in Palermo (35.8%), indicating possible regional variations in torture methods or reporting patterns.

Notably, the medico-legal assessment in Palermo demonstrated a very high rate of consistency between alleged torture and physical findings (over 91% judged as consistent or highly consistent); similarly, the findings of Montpellier emphasized the correlation between specific methods (e.g., burns, FGM) and higher compatibility scores. This supports the utility of structured forensic assessments in corroborating histories of asylum seekers, as recommended by the Istanbul Protocol.

Notable differences were observed between centers. For example, Montpellier reported a significantly lower rate of electrical injuries than that of other centers (1%), while Brescia showed a higher-than-expected frequency of gunshot and electrical trauma, indicating possible variations in regional conflict dynamics, migratory routes, or documentation practices.

This study confirms that the professional's experience is a critical factor in the forensic assessment process. Preferably, these assessments should be conducted in specialized centers with multidisciplinary expertise, capable of maintaining high standards in both clinical and forensic documentation [31,32]. The findings of this study indicate that while agreement was high on injury identification and classification, conversely, greater divergence than that agreement emerged in the interpretive assessment of compatibility, particularly involving the resident in forensic medicine with no prior experience in the assessment of torture sequelae. This supports the hypothesis that experience and specialized training in the forensic assessment of torture significantly enhance the consistency and reliability of evaluative judgments. This study had some limitations that should be acknowledged. First, the limited number of observers, only three, representing different levels of forensic experience, may constrain the generalizability of the inter-rater variability. Moreover, observers assessed the medical records retrospectively without participating in the in-person forensic examination. The lack of direct clinical assessment may have influenced evaluation accuracy, particularly for more nuanced or context-dependent findings. Another limitation is the reliance on a single expert as the reference standard. Lastly, although raters were familiar with the Istanbul Protocol, the absence of a standardized training prior to the evaluation may have contributed to variability in interpretation and application of its criteria.

From a practical perspective, broader and more systematic implementation of the Istanbul Protocol is advisable when clinically indicated, including specialist examinations and imaging techniques to document evidence of violence. A "hub-and-spoke" organizational model could be implemented, in which transit or first-reception centers conduct initial screening and refer suspected cases to specialized facilities for in-depth evaluation [33]. A major challenge in the forensic medical assessment of torture sequelae among asylum seekers is the frequently long interval between the traumatic event and the clinical examination [34]. In many cases, evaluations occur months or years after torture, leaving physical findings often restricted to well-healed scars, which may be challenging to accurately associate with specific forms of violence. In these cases, the use of imaging diagnostic tests can be helpful for the allegation of evidence of torture [35,36]. Minimizing this time gap is crucial to enhance diagnostic accuracy and documentation and facilitate timely and appropriate care [37,38]. Consequently, training of healthcare professionals constitutes another pivotal issue. Both targeted and diffuse educational programs for physicians, psychologists, nurses, and other involved personnel are necessary to enhance the early identification of torture indicators and to ensure timely initiation of protective interventions.

Centralizing these services in qualified facilities enhances health outcomes and supports the determination of international protection, since comprehensive assessment constitutes a key evidentiary component in asylum procedures [5,9]. Adherence to rigorous clinical standards in forensic assessment of torture survivors among asylum seekers is essential to ensure the accuracy, reliability, and ethical integrity in assessments that carry significant legal and humanitarian consequences. A consistent approach across all referral centers, employing common methodologies for physical and psychological examination, consistent classification systems for injuries, and unified criteria for assessing the consequences of trauma, is necessary to ensure the credibility of forensic reports. This standardization enhances the comparability of data across diverse regions and strengthens the global understanding of torture as a widespread human rights violation. From a global humanitarian and public health perspective, the implementation of standardized protocols facilitates more effective advocacy, policy-making, and protective interventions for survivors, while reinforcing the role of forensic medicine in the defense of dignity and justice for vulnerable populations [39–41].

Torture survivors often experience long-term physical and psychological sequelae, potentially evolving into permanent disabilities if not appropriately managed [42]. A comprehensive approach to care is therefore required, encompassing medical, psychological, and social aspects, being multidimensional in nature, and being maintained over time. The intervention must extend beyond the forensic evaluation, encompassing rehabilitation, psychosocial support, and pathways to restore personal integrity [18]. Cultural differences and barriers to

intercultural communication add further complexity to the assessment process. Understanding how survivors perceive and interpret the violence they experienced demands anthropological and psychological perspectives integration [4,43,44]. A person-centered approach that considers the survivor's sociocultural background is therefore crucial. Experience gained from FGM assessment provides a valuable framework, providing insights into sex-based violence and techniques for sensitive interviewing [12,45–48].

There is a critical need to advance research in this area, with investment in studies aimed at identifying and validating clinical, psychological, and forensic markers of torture and ill-treatment. Increased data sharing and exchange of experiences among centers and professionals, at both national and international levels, can support the development of more effective diagnostic tools and the establishment of standardized best practices.

5. Conclusions

In conclusion, this study underscores the central role of specialized forensic expertise in the medico-legal assessment of asylum seekers reporting torture and ill-treatment. Comparative analysis across multiple international centers confirmed the predominance of physical trauma, particularly blunt trauma, as the most commonly documented type of abuse. The inter-center variation in the distribution of injury types likely reflects regional variation in methods of torture and heterogeneity in documentation standards.

Inter-rater reliability was consistently high for both injury identification and classification, as well as for recognizing associated physical and psychological sequelae; however, discrepancies emerged in assigning consistency levels under the Istanbul Protocol, particularly in assessments conducted by less experienced professionals. These findings underscore the necessity of structured training, continuous supervision, and standardized protocols to ensure the accuracy, reproducibility, and evidentiary value of forensic assessments to document human rights violations. Further research with larger samples should be conducted to develop more effective diagnostic tools and to establish standardized practices.

7. Consent to participate

All patients subjected to forensic examination provided informed consent to participate according to national legislation.

Ethics approval

Our investigations were carried out following the rules of the Declaration of Helsinki of 1975, revised in 2013. The research project was approved by the Bioethics Committee of the University of Palermo (Approval n. 137/2023 on 23rd February 2023)

Acknowledgements

This research project was supported under a collaborative protocol involving the Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University of Palermo, and Médecins Sans Frontières (Doctors Without Borders). The authors would like to express their gratitude to Médecins Sans Frontières for their ongoing support of the clinical and forensic activities carried out by the Institute of Forensic Medicine at the University of Palermo.

Data availability

The data that support the findings of the study are available from the corresponding author upon reasonable request.

References

[1] United Nations High Commissioner for Refugees (UNHCR). Global trends: forced displacement in 2020. Geneva: UNHCR; 2020. Available from: https://www.unhcr.org/statistics/unhcrstats/60b638e37/global-trends-forced-displacement-2020.html \[Accessed 21 July 2024].

[2] Ministero dell'Interno. Sbarchi e accoglienza dei migranti: tutti i dati. Available from: http://www.interno.gov.it/it/sala-stampa/dati-e-statistiche/sbarchi-e-accoglienza-dei-migranti-tutti-i-dati \[Accessed 21 July 2024].

[3] Ministero dell'Interno. I numeri dell'asilo. Available from: http://www.libertaciviliimmigrazione.dlci.interno.gov.it/it/documentazione/statistica/i-numeri-dellasilo \[Accessed 24 July 2024].

[4] V. Tullio, C. La Spina, D. Guadagnino, G.D. Albano, S. Zerbo, A. Argo, Ethical and forensic issues in the medico-legal and psychological assessment of women asylum seekers, Healthcare (Basel) 11 (1) (2023) 23, https://doi.org/10.3390/healthcare11010023.

[5] L. Franceschetti, F. Magli, V.G. Merelli, E.A. Muccino, A. Gentilomo, F. Agazzi, et al., The effect of the medico-legal evaluation on asylum seekers in the Metropolitan City of Milan, Italy: a pilot study, Int. J. Legal Med. 133 (3) (2019) 883–891, https://doi.org/10.1007/s00414-019-02003-5.

[6] N.G. Aguirre, A.R. Milewski, J. Shin, D. Ottenheimer, Gender-based violence experienced by women seeking asylum in the United States: a lifetime of multiple traumas inflicted by multiple perpetrators, J. Forensic Leg. Med. 74 (2020) 102026, https://doi.org/10.1016/j.jflm.2020.102026.

[7] G.D. Albano, D. Guadagnino, M. Midiri, C. La Spina, V. Tullio, A. Argo, S. Zerbo, Torture and maltreatment in prison: a medico-legal perspective, Healthcare (Basel) 11 (4) (2023) 576, https://doi.org/10.3390/healthcare11040576.

[8] J. Moffat, S. Tambuzzi, L. Franceschetti, F. Magli, C. Cattaneo, The migrant experience through the lens of forensic medicine: the multifaceted utility of asylum applicants' medico-legal narratives and forensic evidence analysed over time, Int. J. Legal Med. 139 (5) (2025) 2605–2620, https://doi.org/10.1007/s00414-025-03531-0.

[9] R. Aarts, L. van Wanrooij, E. Bloemen, G. Smith, Expert medico-legal reports: the relationship between levels of consistency and judicial outcomes in asylum seekers in the Netherlands, Torture 29 (1) (2019) 36–46, https://doi.org/10.7146/torture.v29i1.111205.

[10] K. Hampton, R. Mishori, Forensic Asylum Evaluation Expert Group. What constitutes a high-quality, comprehensive medico-legal asylum affidavit in the United States immigration context? A multi-sectoral consensus-building modified Delphi, J. Forensic Leg. Med. 96 (2023) 102513, https://doi.org/10.1016/j.jflm.2023.102513.

[11] S. Visentin, G. Pelletti, T. Bajanowski, S.D. Ferrara, Methodology for the identification of vulnerable asylum seekers, Int. J. Legal Med. 131 (6) (2017) 1719–1730, https://doi.org/10.1007/s00414-017-1648-0.

[12] S. Sablone, C.P. Campobasso, C. Tettamanti, A.N. Cardinale, A. Argo, V. Tambone, et al., Gender-based approach: an essential tool in patient safety management and forensic medicine practice, J. Sex Gend. Specif. Med. 10 (3) (2024) 167–168.

[13] A. Saadi, K. Hampton, M.V. de Assis, R. Mishori, H. Habbach, R.J. Haar, Associations between memory loss and trauma in US asylum seekers: a retrospective review of medico-legal affidavits, PLoS One 16 (3) (2021) e0247033, https://doi.org/10.1371/journal.pone.0247033.

[14] G. De Jesús-Rentas, J. Boehnlein, L. Sparr, Central American victims of gang violence as asylum seekers: the role of the forensic expert, J. Am. Acad. Psychiatry Law 38 (4) (2010) 490–498.

[15] A. Sadana, S.E. Gérard, A. Tang, L. Pardee, O. Sher, A. Foote, et al., Medical evidence in asylum applications: medical versus legal approaches, J. Forensic Leg. Med. 97 (2023) 102553, https://doi.org/10.1016/j.jflm.2023.102553.

[16] A.C. Williams, C.R. Peña, A.S. Rice, Persistent pain in survivors of torture: a cohort study, J. Pain Symptom. Manage 40 (5) (2010) 715–722, https://doi.org/10.1016/j.jpainsymman.2010.02.018.

[17] I. Bianchi, M. Focardi, V. Bugelli, F. Pradella, C. Giolli, F. Friani, et al., Tortures alleged by migrants in Italy: compatibility and other medicolegal challenges, Int. J. Legal Med. 135 (6) (2021) 2489–2499, https://doi.org/10.1007/s00414-021-02616-3.

[18] G.F. Longstreth, C. Attix, J. Kuck, Torture survivors and asylum: legal, medical, and psychological perspectives, Am. J. Med. 136 (3) (2023) 244–251, https://doi.org/10.1016/j.amjmed.2022.10.014.

[19] P.R. Cohen, Worldwide forensic dermatology, Cureus 17 (5) (2025) e84645, https://doi.org/10.7759/cureus.84645.

[20] H. Furtmayr, A. Frewer, Documentation of torture and the Istanbul Protocol: applied medical ethics, Med. Health Care Philos. 13 (3) (2010) 279–286, https://doi.org/10.1007/s11019-010-9248-0.

[21] F.A. Akar, R. Arbel, Z. Benninga, M.A. Dia, B. Steiner-Birmanns, The Istanbul Protocol: implementation and education in Israel, Isr. Med. Assoc. J. 16 (3) (2014) 137–141.

[22] United Nations Human Rights (UNHR). Istanbul Protocol: manual on the effective investigation and documentation of torture and other cruel, inhuman or degrading treatment or punishment. Geneva: Office of the High Commissioner of Human Rights; 2022. Available from: [Doi: 10.2307/j.ctt207g8gw](Doi: 10.2307/j.ctt207g8gw).

[23] Médecins Sans Frontières. Inhumane: torture along the Mediterranean migration route, and the support of survivors in a fragile system. 2025. Available from: [https

://prezly.msf.org.uk/torture-along-the-mediterranean-migration-routeand-the-support-of-survivors-in-a-fragile-system](https://prezly.msf.org.uk/torture-along-the-mediterranean-migration-routeand-the-support-of-survivors-in-a-fragile-system) \[Accessed 4 August 2025].

[24] M.C. Russo, E. Azzalini, C. Rossetti, F. Maghin, A. Antonietti, A. Verzeletti, Asylum seekers: forensic remarks regarding 185 cases, Med. Sci. Law 60 (1) (2020) 30–36, https://doi.org/10.1177/0025802419850265.

[25] S. Papadodima, Q. Rebillard, C. Villard, P.A. Peyron, F. Herman, M. Tsellou, et al., Medicolegal evaluation of asylum seekers alleging torture and ill-treatment: a 4-year retrospective study in Montpellier, France, Forensic Sci. Int. 367 (2025) 112357, https://doi.org/10.1016/j.forsciint.2024.112357.

[26] R. Clément, D. Lebossé, L. Barrios, O. Rodat, Asylum seekers alleging torture in their countries: evaluation of a French center, J. Forensic Leg. Med. 46 (2017) 24–29, https://doi.org/10.1016/j.jflm.2016.12.011.

[27] P. Perera, Physical methods of torture and their sequelae: a Sri Lankan perspective, J. Forensic Leg. Med. 14 (3) (2007) 146–150, https://doi.org/10.1016/j.jcfm.2006.05.010.

[28] L. Franceschetti, F. Magli, C. Nava, D. De Angelis, A. Verzeletti, C. Cattaneo, Physical assessment of tortures: interobserver analysis of medico-legal degrees of consistency based on the Istanbul Protocol, Legal Med. 62 (2023) 102248, https://doi.org/10.1016/j.legalmed.2023.102248.

[29] M. Marrone, E. Paladini, F. Pititto, G. Pulin, F. Vinci, A. Marzullo, Histopathological analysis of fuel burns in deceased migrants, J. Travel Med. 32 (1) (2025) taae104, https://doi.org/10.1093/jtm/taae104.

[30] G. Pulin, F. Pititto, R. Bellacicco, E. Paladini, S. Lonero Baldassarra, L. Aventaggiato, et al., It's not always the sea's fault: hydrocarbon poisoning killed two migrants in the central Mediterranean, Clin. Ter. 176 (3) (2025) 266–270, https://doi.org/10.7417/CT.2025.5220.

[31] P. Pérez-Sales, R. González-Rubio, B. Mellor-Marsá, G. Martínez-Alés, Beyond torture checklists: reliability and construct validity of the Torturing Environment Scale (TES), BMC Public Health 21 (2021) 372, https://doi.org/10.1186/s12889-021-10384-w.

[32] C. Cattaneo, S. Tambuzzi, S. De Vecchi, L. Maggioni, G. Costantino, Consequences of the lack of clinical forensic medicine in emergency departments, Int. J. Legal Med. 138 (1) (2024) 139–150, https://doi.org/10.1007/s00414-023-02973-8.

[33] M. Jühling, L.M. König, H. Gruber, V. Wolf, S. Ritz-Timme, F. Mayer, Impact of (forensic) expert opinions according to the Istanbul Protocol in Germany: results of the in\:Fo-project, Int. J. Legal Med. 137 (3) (2023) 863–873, https://doi.org/10.1007/s00414-023-02950-1.

[34] C. Cattaneo, L. Franceschetti, Commentary on: “Impact of (forensic) expert opinions according to the Istanbul Protocol in Germany”, Int. J. Legal Med. 137 (6) (2023) 1939–1940, https://doi.org/10.1007/s00414-023-03063-5.

[35] Albano GD, Re GL, Salerno S, Triskari MB, Fornasari M, Micci G, et al. The role of forensic imaging in the allegations of torture in asylum seekers. Int. J. Legal Med. 2025 (Epub ahead of print). Doi: 10.1007/s00414-025-03601-3.

[36] A. Viero, Z. Obertová, A. Cappella, C. Messina, L.M. Sconfienza, F. Sardanelli, et al., The problem of dating fractures: a retrospective observational study of radiologic features of fracture healing in adults, Forensic Sci. Int. 329 (2021) 111058.

[37] H.D. Petersen, B. Morentin, Assessing the level of credibility of allegations of physical torture, Forensic Sci. Int. 301 (2019) 263–270, https://doi.org/10.1016/j.forsciint.2019.05.043.

[38] K. Clarysse, M. Grosber, J. Ring, J. Gutermuth, C. Kivlahan, Skin lesions, differential diagnosis and practical approach to potential survivors of torture, J. Eur. Acad. Dermatol. Venereol. 33 (7) (2019) 1232–1240, https://doi.org/10.1111/jdv.15439.

[39] R. Cecchi, V. Masotti, M. Sassani, A. Sannella, G. Agugiaro, T. Ikeda, et al., Femicide and forensic pathology: proposal for a shared medico-legal methodology, Legal Med. 60 (2023) 102170, https://doi.org/10.1016/j.legalmed.2023.102170.

[40] J.L. Thomsen, J. Voigt;, Forensic medicine and human rights, Forensic Sci. Int. 36 (1–2) (1988) 147–151, https://doi.org/10.1016/0379-0738(88)90228-9.

[41] I. Ketsekioulafis, K. Katsos, C. Kolentinis, D. Kouzos, K. Moraitis, C. Spiliopoulou, et al., Humanitarian forensic medicine: a systematic review, Int. J. Legal Med. 139 (2) (2025) 747–761, https://doi.org/10.1007/s00414-024-03381-2.

[42] S. Woldu, M. Brasholt, Suspension torture and its physical sequelae, J. Forensic Leg. Med. 80 (2021) 102155, https://doi.org/10.1016/j.jflm.2021.102155.

[43] M. Marrone, F. Pititto, E. Paladini, et al., Bad news, bad behavior: the communication of bad news and aggressive behavior in healthcare. A multicenter retrospective Italian study, J. Patient Saf. Risk Manag. (2025), https://doi.org/10.1177/25160435251366614 in press.

[44] S. Sablone, M. Bellino, A.N. Cardinale, M.G. Violante, R. Risola, A. Magno, et al., Psychological autopsy: a powerful tool in forensic investigations, Forensic Sci. 4 (2024) 635–646, https://doi.org/10.3390/forensicsci4040044.

[45] H. Lever, D. Ottenheimer, J. Teysir, E. Singer, H.G. Atkinson, Depression, anxiety, PTSD and GBV in asylum-seeking women with FGM/C: a retrospective case review, J. Immigr. Minor. Health 21 (3) (2019) 483–489, https://doi.org/10.1007/s10903-018-0782-x.

[46] G. Di Donna, P. Di Lorenzo, C.I. Aquino, M. Marisei, C. Casella, D. Surico, et al., Gender violence during the three ages of life and the impact of the COVID-19 pandemic: a review, Int. J. Soc. Determinants Health Serv. 54 (4) (2024) 423–435.

[47] H. Lever, K.A. Baranowski, D. Ottenheimer, H.G. Atkinson, E.K. Singer, Histories of pervasive GBV in asylum-seeking women with FGM/C, J. Trauma. Stress 35 (3) (2022) 839–851, https://doi.org/10.1002/jts.22791.

[48] R. Mishori, D. Ottenheimer, E. Morris, Conducting an asylum evaluation focused on FGM/C status or risk, Int. J. Gynaecol. Obstet. 153 (1) (2021) 3–10, https://doi.org/10.1002/ijgo.13428.