数字时代下社会危险性的认定方法

尤利亚

西藏民族大学法学院　陕西咸阳

【摘要】社会危险性审查是逮捕决定的核心要素，在人工智能时代，基于算法模型的评估可显著提升结论的准确性与精确度。当人工智能算法得出结论后，对无异议的结论检察机关应当依据结论作出判断；对有异议的结论检察机关应当综合考量风险因素以及反驳意见后作出判断。

【关键词】逮捕；社会危险性；风险评估；推定

【收稿日期】2025 年 5 月 13 日　　【出刊日期】2025 年 6 月 10 日　　【DOI】10.12208/j.ssr.20250238

Methods for identifying social danger in the digital age

Liya You

School of Law, Tibet University for Nationalities, Xianyang, Shaanxi

【**Abstract**】Social danger review is the core element of arrest decisions. In the era of artificial intelligence, evaluation based on algorithm models can significantly improve the accuracy and precision of conclusions. When the artificial intelligence algorithm reaches a conclusion, the procuratorate should make a judgment based on the conclusion for undisputed conclusions; for controversial conclusions, the procuratorate should make a judgment after comprehensively considering risk factors and rebuttal opinions.

【**Keywords**】Arrest; Social danger; Risk assessment; Presumption

引言

社会危险性是逮捕三要件中最为关键、最为难以把握的要件。在实践中多出现“构罪即捕”的现象，只要满足“有证据证明有犯罪事实”以及“可能判处徒刑以上刑罚”检察院就批准逮捕[1]。究其原因，社会危险性无法发挥限制逮捕的作用，没有贯彻比例性原则在强制措施中的使用。为了发挥该要件的作用，我国诸多学者提出解决问题的核心是要构建社会危险性的证明机制。但构建该证明机制有一个前提条件：认定社会危险性是否属于证明？这一问题的争议之处在于认定社会危险性行为之性质理论上存在证明、推定以及评估三种认定方式。

1 证明：难以认定未来事实

在诉讼中，认定社会危险性对应的是未来尚未发生的事实，很难由证据予以证明，从而在论述如何认定社会危险性时普遍采取了推定或是评估方法。少部分学者则认为社会危险性可以予以证明，并且可以通过直接证据予以证明，亦可以通过间接证据予以证明。在探索我国认定社会危险性适合采取推定还是评估前，还是有必要先评述少部分支持证明的学者，考察社会危险性能否如过去事实一样直接予以证明。

1.1 直接证据的转化证明

我国有少部分学者认为社会危险性可以直接证明，支持者的观点缺陷在于“欲意打击报复”这一类犯意表示行为并不能将对社会危险性的证明转化为对预备行为的证据，支持者的观点本质上就是混淆了社会危害性与社会危险性。因此，在犯意表示行为并不能直接转化为预备行为的情况下，仅仅证明叫嚣行为等行为无法直接认定犯罪嫌疑人有社会危险性。

1.2 间接证据的推定证明

除认为社会危险性可以通过直接证据予以转化证明外，有学者认为可以采用推定的方法进行间接证明社会危险性。但推定并非间接证明的证明方法。尽管我国部分学者认为应当将事实推定归类为纯粹的间接证明；但是有更多的学者指出，如果从基础事实（间接性事实）到推定事实（待证事实）的认定上分析，事实推

作者简介：尤利亚（1997-）女，汉族，河南新乡人，西藏民族大学法学院硕士研究生在读，研究方向：不区分研究方向。

定就不同于间接证明。

事实推定和间接证明都提及经验法则的运用，在事实推定中，基础事实到推定事实的过程需要经验法则的运用；间接证明中诸多间接事实之间相互印证需要经验法则发现其内在逻辑联系。尽管事实推定与间接证明均要求“经验法则”，但经验程度要求不同，间接证据主要运用于发现各间接证据的逻辑联系，经验盖然性程度低；事实推定要求经验法则作为推定大前提的存在，因而需要具有较高盖然性。

当推定不属于证明时，纯粹间接证据无法证明社会危险性，支持间接证明的理论即丧失了理论基础。但这并不能直接否定间接证明无法认定社会危险性，间接证明所采取的证明方法无法证明社会危险性才是否定社会危险性无法证明的前提。实质上，影响社会危险性的考虑因素较多，年龄、性别、犯罪者的世界观和政治思想、知识和道德水平、人格、心理特点、犯罪者的犯罪原因、社会经历、累犯等因素都会影响社会危险性的认定。

综上所述，间接证明基于社会危险性本身的性质，难以以认定过去事实之证明方法认定未来事实存在的可能性。唯有推定以及评估两种方法可以跨越基础事实到待证事实的“鸿沟”。

2 推定：无法兼容否定因素

在普遍承认无法直接证明社会危险性的前提下，部分学者主张可以通过证明基础事实以推定存在社会危险性。支持推定的学者一般都认为社会危险性推定为法律推定。这是因为事实推定的正当性欠缺，若是允许事实推定，则法官可以任意行使自由裁量权，随意转移证明责任。即便是少部分支持事实推定的学者，也因为事实推定的正当性问题，将事实推定隐藏在间接证明中[2]。在多数学者支持社会危险性认定为法律推定的前提下，我们首先需要分析法律推定的形式要求，以确定《刑事诉讼法》对社会危险性认定之规定是否属于法律推定。

2.1 法律推定的形式标准

我国学者支持社会危险性认定属于法律推定的原因为：《刑事诉讼法》第 81 条规定了存在社会危险性的五种情形以及《人民检察院刑事诉讼规则》第 129 条到 133 条又细化了《刑事诉讼法》规定的五种情形。这导致学者们认为《刑事诉讼法》规定了社会危险性的基础情形，只要证明这五种情形存在一种即可推定存在社会危险性。但是，社会危险性的认定可以界定为法律推定吗？

作为确立在文本中的推定，法律推定因为转移了证明责任，必须有其明确的形式要求。例如，在刑法领域，持有型犯罪是否属于法律推定存在较大争议。反对者认为《刑法》已经规定了持有型犯罪完备的构成要件，并没有确定推定规范；支持者认为检察机关只要证明了被告人持有某种违禁品，即可推定存在明知，进而法院可以直接认定存在持有违禁品这一推定事实的成立。

据此，我们可以总结出法律推定的形式要件：基础事实+（证明责任转移）=推定事实。以此形式规范套用到持有型犯罪之后，持有型犯罪不符合法律推定的形式要件，缺少了证明责任的转移。以持有型犯罪为参照，法律推定必须通过法律规范明确证明责任转移，而非可以通过经验法则或是逻辑法则推定出证明责任转移。

2.2 社会危险性之法律推定形式要件的缺位

在分析法律推定的形式要件后，我们总结出法律推定必须有法律规范明确规定基础事实以及证明责任转移。将这一形式规范用于社会危险性的认定中，但《刑事诉讼法》第 81 条的规范在基础事实与证明责任转移的规定上不符合法律推定的形式要求。该条款为基础事实规定了其他可能情形。尽管从表面上看，该款的五项规定可以作为认定社会危险性的基础事实，只要证明存在这五种情形就可以认定存在社会危险性。但是，《人民检察院刑事诉讼规则》在细化五种情形时，又在每一种情形最后增加了“其他可能……的情形”这一项，导致依据法律规定推定社会危险性是不周延的。

2.3 社会危险性认定无法适用事实推定

当社会危险性认定不符合法律推定的形式要求时，检察院或法院将无法依据《刑事诉讼法》第 81 条第 1 款的五项规定直接认定是否存在社会危险性。既然否定了法律推定，社会危险性认定可以采取事实推定吗？

事实推定违背了无罪推定原则。传统观点认为，在推定中，如果证实了基础事实，证明推定事实不存在的举证责任就转换给被告人。如果将事实推定视为推定，那么法官就可以随意转移证明责任，这无疑违背了无罪推定原则[3]。因此，如何理解证明责任的转移是需要研究的问题。社会危险性认定不符合推定的形式要件。即便部分学者承认事实推定的存在是为了解决“事实真伪不明”的疑案而不得不进行的“理性妥协”，而忽视了事实推定的正当性基础，我们不能将社会危险性认定理解为事实推定[4]。

因此，《刑事诉讼法》第 81 条第 1 款及第 2 款的规定均不符合推定的形式要求，尤其是第 2 款规定的考虑因素所产生的否定因素导致了实践中难以推定社

会危险性。社会危险性的认定必然需要将所有的肯定因素及否定因素综合考虑，才能确保其准确性，这不符合推定证明责任转移以及单一基础事实的形式要求。

3 评估：得出结论后的证明难题

鉴于综合考虑认定社会危险性之肯定因素和否定因素的主观性，尤其是证明与推定均难以综合考虑肯定因素和否定因素的情况下，多数学者指出需要引入具备客观性、规范性的评估机制，使最终的结果更为标准化和统一化。

3.1 目前我国采取评估方式的进路有两种

第三方风险评估机制与人工智能算法评估机制。这两者评估方式的性质不同，前者属于临床评估；后者属于精算评估。

以上两种精算评估和结构化临床评估为主，尽管精算评估与结构化临床评估因群体预测存在歧视风险，相较于临床评估却能够实现风险评估的准确性和精确性[5]。因此即便是在美国，审前服务官使用的传统的临床评估也逐渐被人工智能算法评估所取代。既然第三方评估都已经在逐步使用算法评估，我国在尚未完全采取第三方评估的前提下，直接由检察院和法院组织研发人工智能算法评估更为适宜。

3.2 如何对待评估结论

评估工具的选择并非使用评估方法的困境，大陆法系国家普遍使用第三方评估，而英国和美国则在大力开发人工智能算法评估，每一种评估方式都有契合社会危险性的方面。但无论是第三方评估还是算法评估，作出判断的并非专业的司法人员，如何认定评估结论是使用评估认定社会危险性的难题。

我国学者亦尚未厘清评估与证明的关系，以至于陷入评估结论需要证明的自相矛盾中，甚至脱离了评估结论的性质而仅仅研究评估方法。对于评估和证明的关系问题，我国部分学者认为评估仅仅作为一种手段，其结论的准确性依然需要证明；又有部分学者认为除有与评估结论相反的充分理由外，检察机关可以直接依据评估结论认定是否存在社会危险性。这两种观点均存在缺陷：

支持对评估结论进行证明的学者忽视了如何证明的问题。无论是第三方评估还是算法评估，在评估时已经考虑了社会危险性各因素，若侦查机关及犯罪嫌疑人对风险评估结论无异议，当检察官作出的判断无法识别是以评估为依据对社会危险性进行证明，还是直接依据评估结果作出判断。

对于算法评估而言，直接依据评估结论认定社会危险性则会产生伦理问题。如美国的 COMPAS 被诟病的原因在于，即便在预测累犯时没有使用罪犯的种族，但它似乎还是提供了有种族差异的结果[6]。这表明算法可能存在歧视，若完全以评估结论作出判断，一方面会侵犯检察机关的司法权；另一方面算法黑箱性可能会忽视个体差异而从某一群体的特征带来歧视和历史偏见。学者们认为社会危险性认定使用推定方法违背了法律面前人人平等原则以及比例原则，又认为使用评估方法忽视了个体间的差异，难以在认定社会危险性的方法中找到平衡。

综上所述，评估是最为复杂的一种认定社会危险性的方式，从形式上可以分为临床评估、精算评估以及结构化临床评估；从方法上分为第三方评估和人工智能算法评估。我国正处于评估体系的初步研发阶段，相较于第三方评估，人工智能评估更具客观性及中立性，采取人工智能算法评估下的结构化临床评估较为适宜。

4 结论

认定社会危险性是批准逮捕的核心，亦是采取其他强制措施的决定性因素，在选择认定社会危险性方法时需要兼顾考虑是否适合采取其他强制措施。证明与推定都只能基于仅有肯定因素的情况下认定社会危险性，唯有评估才能优先考虑否定因素以及其他强制措施[7]。未来认定社会危险性的问题核心不应当为认定方法问题，而是如何设计人工智能算法评估体系。目前我国仍处于设计社会危险性量化评估模型的初级阶段，各省选择的量化指标不够统一、司法人员对评估结果信任度不足、人工智能技术存在不当应用等问题依然突出。如何在轻罪治理背景下进一步推定社会危险性评估之准确性、科学性，尚需理论界与实务界更多的关注与研究。

参考文献

[1] 参见王鹏飞:《普遍处罚模式下预备行为的类型建构》,《法学杂志》2021 年第 6 期,第 121 页.

[2] 参见杨依:《以社会危险性审查为核心的逮捕条件重构——基于经验事实的理论反思》,《比较法研究》2018 年第 3 期,第 138 页.

[3] 参见褚福民:《事实推定的客观存在及其正当性质疑》,《中外法学》2010 年第 5 期,第 669-670 页.

[4] 龙宗智:《“印证”的治理》,《法学家》2022 年第 2 期,第 101 页.

[5] 孙谦:《司法改革背景下逮捕的若干问题研究》,《中国法学》2017 年第 3 期,第 38 页.

[6] 胡印富、David Field:《澳大利亚刑事推定证据规则适用研究》,《河北法学》2015 年第 4 期,第 106 页.

[7] 裴苍龄:《再论推定》,《法学研究》2006 年第 3 期,第 120 页.

版权声明：©2025 作者与开放获取期刊研究中心（OAJRC）所有。本文章按照知识共享署名许可条款发表。
http://creativecommons.org/licenses/by/4.0/

OPEN ACCESS