Multidisciplinary Output Research For Actual and International Issue (MORFAI Journal) ISSN (e): 2808-6635
Volumes 6 No. 6 (2026)

WELFARE RANKINGS AND SOCIAL ASSISTANCE ALLOCATION: A SYSTEMATIC REVIEW WITH IMPLICATIONS FOR INDONESIA

Muh Alfian[1]*, Hetifah Sjaifudian[2], Muhaidir Ikram[3] Nuzul Husnah[4]
[1]Universitas Muhammadiyah Jakarta, Indonesia
[2,4]Politeknik Jatiluhur, Indonesia
[3]Program Studi Manajemen, Universitas Negeri Makassar, Indonesia
E-mail: muhammad.alfian38@gmail.com[1]*, hetifah@polijati.ac.id[2], muhaidir.ikram@unm.ac.id[3], nuzulhusnah@polijati.ac.id[4]

Received:05/10/2026 | **Revised** : 12/10/2026 | **Accepted:** 30/10/2026 | **Published:** 24/11/2026

Abstract

Welfare rankings are widely used to identify disadvantaged households and prioritize social assistance, yet accurate socioeconomic classifications do not necessarily ensure equitable or responsive benefit allocation. This systematic literature review examines the relationship between welfare classification, targeting accuracy, distributive equity, dynamic vulnerability, and institutional governance, with particular attention to implications for Indonesia. Literature published between 2017 and 2026 was identified through Scopus and Google Scholar using Publish or Perish. Following screening, methodological appraisal, and evidence-use adjudication, a provisional set of 18 studies was synthesized using a thematic narrative approach. The findings indicate that targeting performance depends on the welfare indicators, evaluation benchmarks, and programme objectives used to define eligibility. Methods that perform relatively well in identifying persistent poverty may be less effective in capturing seasonal deprivation or emerging vulnerability. Furthermore, socioeconomic ranking accuracy does not necessarily translate into equitable allocation because administrative procedures, eligibility requirements, local discretion, and institutional incentives influence actual beneficiary selection. Static classifications may also become less representative as household circumstances change. Based on these findings, the review develops a six-stage conceptual framework linking welfare-need definition, measurement, eligibility determination, allocation, reassessment, and performance evaluation. For Indonesia, the synthesis highlights the importance of evaluating DTSEN welfare deciles against independently measured deprivation, actual benefit receipt, and responsiveness to changing vulnerability. The review contributes an integrated perspective on targeting while emphasizing the need for rigorous empirical evaluation of welfare classification and allocation reforms.

Keywords: Welfare Rankings; Social Assistance Allocation; Targeting Accuracy; Distributive Equity; Dynamic Vulnerability.

INTRODUCTION

Social assistance plays a central role in protecting economically vulnerable populations, reducing poverty, and strengthening households' capacity to cope with socioeconomic disruptions. Despite the expansion of social protection systems, substantial coverage gaps remain. Approximately 52.4% of the global population was covered by at least one social protection benefit in 2024, while around two billion people in low- and middle-income countries remained uncovered or inadequately protected (International Labour Organization, 2024; World Bank, 2025). These gaps raise important questions concerning not only the adequacy and availability of social assistance but also the effectiveness of mechanisms used to identify vulnerable populations and distribute limited public resources. In countries where household income is difficult to verify and administrative capacity remains constrained, determining who should receive assistance represents a complex allocation problem involving efficiency, equity, and social welfare considerations. Decisions between targeted and universal transfers involve important redistributional trade-offs, while even relatively sophisticated econometric targeting methods may fail to reach substantial proportions of the poorest households (Brown et al., 2018; Hanna & Olken, 2018). Beneficiary identification, redistribution, and implementation also represent interconnected challenges in designing effective social protection systems in developing economies (Banerjee et al., 2024). These findings underscore the importance of examining whether welfare classification systems translate into meaningful improvements in social assistance allocation rather than assuming that the identification of disadvantaged households necessarily produces equitable distributional outcomes.

Welfare rankings are commonly employed to prioritize households according to estimated socioeconomic conditions. Approaches such as proxy means testing (PMT), community-based targeting (CBT), categorical eligibility, geographic targeting, and hybrid statistical methods differ in their assumptions, information requirements, and administrative implications. However, the use of welfare rankings introduces an important distinction between estimating household disadvantage and determining actual programme eligibility. Evidence from nine African countries demonstrates that conventional PMT approaches can exclude substantial numbers of poor households despite their ability to distinguish between poorer and wealthier populations (Brown et al., 2018). Similarly, comparative evidence from nine social safety net programmes in the Sahel indicates that the relative performance of PMT and CBT depends on targeting objectives, beneficiary shares, and the welfare indicators used for evaluation (Schnitzer & Stoeffler, 2024). Evidence from Indonesia provides an additional perspective on the institutional conditions affecting welfare-based targeting. The coverage of household registration is an important determinant of targeting effectiveness, indicating that improvements in socioeconomic ranking alone may be insufficient when vulnerable households are absent from the underlying database (Bah et al., 2019). Unified targeting arrangements can also improve targeting performance and support complementarities across social programmes, highlighting the importance of examining beneficiary identification within a broader social assistance system (Tohari et al., 2019). Comparative research further demonstrates that alternative targeting methods may produce relatively similar outcomes under certain evaluation criteria, making administrative costs and implementation arrangements relevant considerations when selecting targeting approaches (Karlan & Thuysbaert, 2019). Experimental evidence from Niger additionally indicates that targeting methods differ not only in their efficiency but also in their perceived legitimacy and subsequent programme impacts (Premand & Schnitzer, 2021). Meanwhile, statistical indicators and community preferences can be combined within a hybrid targeting framework, offering an alternative to relying exclusively on either statistical or community-based classification (Follett & Henderson, 2023). Together, these studies suggest that the effectiveness of welfare rankings cannot be assessed independently of registration coverage, programme objectives, eligibility rules, and the conditions under which classifications are translated into beneficiary selection.

A further challenge concerns the conceptualization and measurement of welfare itself. Conventional targeting approaches frequently rely on income, expenditure, or observable household assets as proxies for economic deprivation. Yet these indicators may not fully capture the multidimensional characteristics of disadvantage or the particular forms of vulnerability that social assistance programmes seek to address. Targeting based on human capabilities can identify a substantially different beneficiary population from approaches emphasizing conventional economic indicators, suggesting that alternative definitions of welfare may generate different socioeconomic rankings even when the same population is assessed (Henderson & Follett, 2022). The apparent performance of targeting methods is also sensitive to the benchmark against which classification errors are evaluated. Biases in the data used to construct welfare predictions can generate unequal welfare losses that are not adequately reflected in conventional accuracy measures, particularly when the severity of deprivation among incorrectly classified households is considered (Dietrich et al., 2024). Accordingly, improvements in predictive performance should not automatically be interpreted as improvements in distributive equity or social welfare. A further distinction arises between targeting households experiencing the greatest deprivation and targeting those expected to benefit most from an intervention, demonstrating that beneficiary prioritization involves distributional objectives that may not be captured by conventional poverty rankings alone (Haushofer et al., 2025). The increasing availability of alternative data sources introduces additional complexities. Machine-learning models combining mobile phone and household survey information can improve the identification of extremely poor households compared with approaches relying on individual data sources, although differences in access to digital information may influence which households are represented in such models (Aiken et al., 2023). Evaluations therefore need to consider not only how many households are classified correctly but also which households are excluded, the severity of their deprivation, and the consequences of allocation errors. More broadly, the choice of welfare indicators, targeting thresholds, and evaluation metrics should be understood as part of the policy decision process rather than as exclusively technical features of statistical classification.

The distinction between statistical classification and actual programme outcomes becomes particularly important when administrative implementation is considered. Beneficiary allocation is influenced by eligibility rules, local discretion, institutional incentives, geographic priorities, and the effectiveness of delivery mechanisms. Political patronage and local institutional arrangements can influence beneficiary selection within Ghana's Livelihood Empowerment Against Poverty programme, while the implementation of China's rural Dibao programme may diverge from its formal poverty-alleviation objectives (Akyeampong, 2024; Li & Walker, 2018). Further evidence from China identifies tensions between centralized administrative requirements and local discretion in

OPEN ACCESS

social assistance implementation, illustrating how institutional arrangements can influence the translation of eligibility criteria into actual programme decisions (Li & Walker, 2021). Nevertheless, decentralization does not necessarily produce uniformly adverse targeting outcomes. Evidence from Malawi demonstrates that local leaders may consider expected productive returns alongside household poverty, creating differences between poverty-based targeting and productive efficiency (Basurto et al., 2020). Experimental evidence from Ethiopia similarly indicates that community leaders' allocation decisions respond to budgetary constraints and the degree of discretion permitted within targeting arrangements (Abay et al., 2026). Evidence from Indonesia further demonstrates that programme delivery arrangements can substantially influence the distribution of assistance even when the intended beneficiary population has already been identified. Replacing in-kind food distribution with electronic vouchers increased the amount of assistance received by targeted households and reduced poverty among the poorest households in a large-scale randomized evaluation, highlighting the distinction between beneficiary identification and effective benefit delivery (Banerjee et al., 2023). Geographic analyses provide an additional perspective, suggesting that the spatial distribution of assistance may reflect particular normative assumptions concerning socioeconomic disadvantage rather than corresponding directly to all dimensions of social need (Poirier, 2020). Taken together, these studies indicate that targeting effectiveness depends on interactions among welfare classification, administrative decision-making, institutional incentives, and programme implementation rather than on statistical identification alone.

The temporal dimension of welfare classification presents another challenge. Household circumstances are not static, and classifications constructed from information collected at a particular point in time may become less representative as households experience income fluctuations, employment changes, migration, or environmental shocks. Although PMT performed better than CBT in identifying consumption-poor households at the time of targeting in Burkina Faso, this advantage disappeared when performance was evaluated against household poverty status 12 and 30 months later (Hillebrecht et al., 2023). This finding suggests that a method's predictive value at baseline may not reflect its effectiveness when programme benefits are ultimately delivered. Longitudinal evidence from Mexico's conditional cash transfer programme further demonstrates the relevance of poverty transitions for understanding beneficiary graduation and programme design (Villa & Niño-Zarazúa, 2019). Institutional investigations also indicate that outdated socioeconomic information and administrative constraints can undermine the responsiveness of social protection systems, as illustrated by the Philippines' Pantawid programme and its associated social registry (Dadap-Cantal et al., 2021). The importance of adaptive targeting becomes particularly evident during crises. Machine learning and mobile phone data can support humanitarian assistance targeting when conventional administrative information is unavailable or outdated, although these approaches also introduce risks of exclusion associated with phone ownership, registration, and data availability (Aiken et al., 2022). Further evidence from Colombia demonstrates that the distributional consequences of static and dynamic targeting approaches depend on household income fluctuations, eligibility thresholds, economic conditions, and policy objectives (Beuermann et al., 2025). Importantly, updating socioeconomic information does not necessarily improve social welfare unless the updated information is appropriately incorporated into targeting decisions (Beuermann et al., 2025). Collectively, these studies suggest that adaptive targeting requires more than periodic data collection; it also involves the capacity of administrative systems to reassess eligibility and respond to changes in household vulnerability.

These concerns are especially relevant to Indonesia, where social assistance administration increasingly relies on integrated socioeconomic information and standardized household welfare rankings. The establishment of the Data Tunggal Sosial dan Ekonomi Nasional (DTSEN), supported by Presidential Instruction No. 4 of 2025 and Minister of Social Affairs Regulation No. 3 of 2025, represents an institutional effort to strengthen the consistency and use of socioeconomic information across public programmes. Within this system, welfare deciles represent relative positions in the socioeconomic distribution rather than direct measurements of absolute poverty or automatic entitlement to assistance (Badan Pusat Statistik, 2026). Earlier evidence from Indonesia demonstrates that the effectiveness of centralized social assistance databases depends on both the coverage of household registration and the accuracy of socioeconomic ranking (Bah et al., 2019). Unified targeting arrangements can also improve the performance of multiple social assistance programmes, suggesting that the benefits of data integration should be evaluated across programme interactions rather than through isolated classification outcomes (Tohari et al., 2019). More recent research provides additional evidence concerning the methodological challenges associated with welfare estimation. Comparisons between statistical and machine-learning methods for PMT-based poverty identification across three Indonesian districts demonstrate opportunities to improve targeting predictions through alternative modelling approaches (Wobcke & Mariyah, 2023). The contribution of area-level characteristics to PMT performance also varies across rural, semi-urban, and urban districts, indicating that the usefulness of geographical

OPEN ACCESS

information depends on local characteristics and spatial granularity (Mariyah & Wobcke, 2025a). Further methodological developments involving domain adaptation seek to address changes in data distributions and improve poverty-targeting estimation under different classification conditions (Mariyah & Wobcke, 2025b). These studies illustrate the potential of methodological innovation while emphasizing that predictive performance requires context-specific validation. Nevertheless, improvements in welfare estimation should be distinguished from observed improvements in eligibility determination, beneficiary receipt, and administrative responsiveness. The development of DTSEN therefore provides a policy-relevant context for examining how standardized welfare rankings interact with programme rules and implementation arrangements, without presuming that data integration itself guarantees more equitable assistance allocation.

Existing research has generated substantial evidence on the design and performance of social assistance targeting systems. Empirical studies have examined econometric prediction, community participation, welfare-based evaluation, programme delivery, and dynamic vulnerability, while broader literature reviews have addressed beneficiary identification and the institutional design of social protection programmes (Banerjee et al., 2024). More focused reviews have also investigated important dimensions of targeting and implementation. Administrative, structural, and political factors have been identified as relevant contributors to exclusion errors within Brazil's Bolsa Família programme, illustrating the importance of implementation conditions for beneficiary access (Pontuschka et al., 2025). Research on digital welfare systems in developing countries further demonstrates that digitalization can reshape inclusion, exclusion, and citizen–state relationships, potentially improving administrative coordination while introducing additional accountability challenges (Huang & Oteng, 2026). These contributions demonstrate that administrative exclusion, institutional governance, and the distributional consequences of social protection have already received substantial scholarly attention. Nevertheless, the available evidence also presents methodological and conceptual challenges for cross-study interpretation. Model-validation studies commonly assess predicted welfare or simulated targeting errors, whereas programme evaluations examine actual benefit receipt, geographical analyses focus on spatial distribution, and institutional investigations emphasize governance and implementation practices. Studies of poverty transitions and emergency assistance introduce an additional temporal dimension. Furthermore, targeting performance may be evaluated using consumption poverty, multidimensional deprivation, beneficiary coverage, predictive accuracy, social welfare losses, or programme impacts. Because these approaches employ different analytical units, welfare benchmarks, and outcome measures, their findings cannot necessarily be compared directly or interpreted as estimates of the same underlying targeting performance. The analytical challenge is therefore not the absence of research on these individual dimensions but the need to clarify how different forms of evidence relate to one another and what they collectively reveal about the relationship between welfare rankings and social assistance allocation. Addressing this challenge requires distinguishing classification performance from distributive outcomes while examining how measurement choices, eligibility thresholds, institutional arrangements, and temporal conditions influence the interpretation of targeting effectiveness.

Against this background, this systematic literature review examines the relationships between welfare rankings, beneficiary eligibility, targeting accuracy, distributive equity, and dynamic vulnerability in social assistance allocation. Focusing on literature published between 2017 and 2026, the review seeks to synthesize evidence across different welfare classification approaches, programme settings, and evaluation methods. Rather than assuming that a particular targeting mechanism is universally superior, the review considers how targeting performance varies according to the definition of welfare, the objectives of beneficiary selection, the distinction between simulated and actual programme outcomes, and the institutional and temporal conditions under which assistance is implemented. Accordingly, the study addresses five interconnected research questions:

RQ1. How have welfare rankings and socioeconomic classifications been operationalized for beneficiary targeting in social assistance programs?

RQ2. What trade-offs between targeting accuracy, distributive equity, and allocation efficiency are identified across different welfare-targeting approaches?

RQ3. How do household welfare dynamics, economic shocks, and institutional conditions shape the performance of social assistance targeting?

RQ4. What integrated framework can explain the relationship between welfare classification, targeting decisions, and equitable social assistance allocation?

RQ5. What evidence-based governance and implementation practices can inform more accurate, equitable, and adaptive social assistance targeting systems?

The review seeks to contribute by systematically differentiating actual programme outcomes from simulated and model-based targeting estimates, identifying the conditions associated with convergent and divergent empirical

findings, and organizing the available evidence within an analytical framework linking welfare measurement, eligibility determination, allocation, reassessment, and performance evaluation. Rather than introducing previously unexplored relationships or identifying a universally optimal targeting method, the review aims to clarify how technical, institutional, and temporal considerations jointly shape social assistance allocation. The resulting synthesis is intended to inform future research and provide evidence-informed considerations for social protection governance, with particular attention to Indonesia's welfare classification system and the implementation of DTSEN.

METHOD

Research Design

This study employed a systematic literature review (SLR) to synthesize evidence concerning welfare rankings and social assistance allocation. The review examined socioeconomic classification, beneficiary targeting, targeting accuracy, distributive equity, dynamic household vulnerability, and institutional governance. The review process was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement to support transparent reporting of study identification, screening, eligibility assessment, and inclusion (Page et al., 2021). Search documentation was additionally informed by the PRISMA-S extension (Rethlefsen et al., 2021). Given the heterogeneity of research designs, welfare indicators, targeting mechanisms, and allocation outcomes, a thematic narrative synthesis was adopted to identify recurring patterns, contrasting findings, and contextual explanations across studies. Statistical meta-analysis was not undertaken because the available evidence did not provide sufficiently comparable outcome definitions, analytical units, or effect measures for meaningful quantitative pooling.

Review Framework and Eligibility Criteria

The review was structured using the Population, Intervention or Phenomenon of Interest, Comparison, Outcomes, and Context (PICOC) framework. This framework established the conceptual boundaries of the review and guided the development of eligibility criteria, literature searches, and evidence synthesis. The comparison component was applied flexibly to accommodate qualitative and observational studies without formal comparison groups, provided they offered relevant evidence concerning welfare classification, beneficiary eligibility, or social assistance allocation.

Table 1. PICOC Framework

Component	Description
Population	Poor, low-income, near-poor, and socioeconomically vulnerable individuals or households.
Intervention / Phenomenon of Interest	Welfare rankings and socioeconomic classification approaches used in social assistance targeting, including proxy means testing (PMT), community-based targeting (CBT), geographic targeting, social registries, beneficiary selection, and adaptive targeting mechanisms.
Comparison	Alternative welfare classification methods, beneficiary selection mechanisms, eligibility criteria, and static versus dynamic targeting arrangements.
Outcomes	Targeting accuracy, inclusion and exclusion errors, distributive equity, allocation efficiency, beneficiary coverage, and responsiveness to changing household vulnerability.
Context	Public social assistance and related social protection programmes across geographical and institutional settings.

The eligibility criteria were developed to identify original research examining the relationship between welfare classification and social assistance targeting or allocation. The review considered peer-reviewed empirical and evaluative journal articles published between 2017 and 2026, including quantitative, qualitative, mixed-methods, experimental, longitudinal, and relevant simulation-based investigations. Documentary and policy analyses were considered when they presented original, traceable evidence directly related to beneficiary eligibility, targeting, or allocation. Studies focusing exclusively on general poverty outcomes or programme impacts without examining the identification, eligibility, or allocation of beneficiaries were excluded from the substantive synthesis.

Table 2. Inclusion and Exclusion Criteria

Criterion	Inclusion Criteria	Exclusion Criteria
Publication period	Studies published between 2017 and the applicable search date in 2026.	Studies published outside the specified period.
Publication type	Original peer-reviewed empirical or evaluative journal articles.	Editorials, conference abstracts, book reviews, non-peer-reviewed publications,

OPEN ACCESS

		and publications without relevant original analysis.
Language	English-language publications, subject to confirmation of the restriction in the original screening protocol.	Other-language publications only where a prospective English-language restriction can be documented.
Population	Poor, low-income, near-poor, and socioeconomically vulnerable individuals or households.	Studies without relevant socioeconomic vulnerability or beneficiary populations.
Research focus	Welfare classification, beneficiary targeting, eligibility determination, social assistance allocation, and related governance mechanisms.	Studies unrelated to targeting, eligibility, or assistance allocation.
Research design	Quantitative, qualitative, mixed-methods, experimental, longitudinal, modelling, simulation, and relevant documentary analyses.	Publications lacking sufficient methodological information to support interpretation of relevant findings.
Outcomes	Targeting accuracy, inclusion and exclusion errors, distributive equity, allocation efficiency, beneficiary coverage, dynamic vulnerability, and targeting governance.	Studies without outcomes or substantive findings relevant to beneficiary identification, eligibility, or allocation.

The criteria were applied to the substantive content of each study rather than to database indexing status alone. Publications retrieved through Google Scholar were therefore not automatically excluded because they lacked Scopus indexing. Where publication metadata, document type, or eligibility information remained uncertain, the records were retained for further verification rather than classified as definitive exclusions.

Information Sources and Search Strategy

The literature search was conducted using Scopus and Google Scholar through Publish or Perish (PoP), covering publications from 2017 to 2026. The historical search records were dated October 19, 2026. The search strategy was developed from the PICOC framework using combinations of terms related to *social assistance, social protection, poverty targeting, beneficiary targeting, proxy means testing, welfare ranking, social registry, cash transfers, beneficiary selection, social safety nets, community-based targeting, inclusion errors, exclusion errors, targeting accuracy, targeting effectiveness, distributive equity, allocation efficiency, dynamic targeting, poverty dynamics, economic shocks, adaptive social protection, administrative data, data quality, community verification, eligibility, and governance*. The search programme was initially organized into 38 thematic entries comprising 10 main searches (M01–M10), six supplementary searches (S11–S16), and 22 expanded searches (E17–E38). These entries were intended to capture variations in terminology across the five research questions. The historical retrieval audit documented completed exports for 11 Scopus searches and 11 Google Scholar searches; therefore, the planned 38-entry structure was not treated as evidence that all entries had been executed. Search expressions were adapted to the field configurations of PoP, and the available database-specific queries, retrieval counts, and export information were recorded in the search log. A subsequent expansion of the strategy was developed but was not included in the historical retrieval results because the revised queries had not yet been executed.

The documented searches yielded 68 records from Scopus and 55 from Google Scholar, together with four additional records from an export whose database provenance could not be conclusively established. Accordingly, the consolidated retrieval contained 127 records. After removing nine duplicates, 118 unique records remained for screening. Duplicate identification was based on bibliographic information, particularly DOI and article-title matching. The historical search log was retained to distinguish executed searches from planned or reconstructed queries. Exact Google Scholar query expressions could be recovered from the original export metadata, whereas the complete input syntax of several historical Scopus searches remained unavailable for independent verification. Consequently, the search audit provides a traceable account of the historical retrieval but does not yet establish complete reproducibility of all database queries.

Study Selection and Quality Appraisal

Study selection involved title and abstract screening, full-text eligibility assessment, and methodological quality appraisal. Following deduplication, 118 unique records were screened against the eligibility criteria. Eighty-two records were excluded at the title and abstract stage, while two records required further bibliographic verification. The remaining 34 reports were identified for full-text retrieval and assessment. Of these, 24 underwent substantive full-text examination, seven had not been obtained, and three remained awaiting comprehensive

OPEN ACCESS

assessment. One report was excluded after full-text examination because its aggregate transfer outcomes could not be sufficiently connected to the review questions. The remaining 23 studies were considered provisionally eligible under the PICOC framework and proceeded to methodological appraisal. These figures represent the documented interim selection audit rather than a finalized PRISMA inclusion flow.

Methodological quality was evaluated using criteria appropriate to the research design of each study. Qualitative, nonrandomized quantitative, and mixed-methods studies were assessed using criteria adapted from the Mixed Methods Appraisal Tool (MMAT), version 2018 (Hong et al., 2018). The cluster-randomized study underwent a preliminary domain-level assessment informed by the revised Cochrane Risk of Bias tool (RoB 2) (Sterne et al., 2019), although the full cluster-specific assessment was not completed. Simulation and predictive-modelling studies were evaluated using adapted criteria addressing data provenance, welfare benchmark validity, model specification, validation procedures, uncertainty, and practical applicability. Documentary analyses were assessed through criteria emphasizing source transparency, analytical coherence, and the relevance of reported evidence to targeting and allocation processes. These adapted criteria were not treated as independently validated risk-of-bias instruments.

The appraisal covered 23 studies and generated 124 item-level assessments. Methodological concerns were recorded qualitatively without calculating a common numerical quality score across different research designs. Following appraisal, studies were classified according to their appropriate contribution to the synthesis: core evidence, conditional evidence, or supporting literature. Fourteen studies were classified as core evidence and four as conditional evidence, forming an 18-study working substantive synthesis. Five additional studies were retained as supporting literature because methodological or applicability limitations restricted the claims that could reasonably be derived from them. These evidence-use categories were distinguished from PICOC eligibility decisions and did not automatically constitute grounds for formal study exclusion. Screening and appraisal decisions were documented through a preliminary single-reviewer audit; independent second-reviewer verification had not been completed.

Data Extraction

A structured data extraction matrix was developed to systematically record the characteristics, methods, findings, and methodological limitations of studies retained for substantive synthesis. The extracted information included authorship, publication year, geographical setting, programme context, research design, data sources, sample characteristics, welfare indicators, targeting mechanisms, eligibility criteria, outcome measures, and principal findings. Particular attention was given to identifying the welfare benchmark underlying each targeting result and distinguishing observed programme outcomes from simulated allocation performance and predictive model estimates.

The working extraction matrix comprised 18 study-level records and 36 claim-level evidence entries. Claim-level extraction documented the relevant source, substantive findings, methodological qualifications, and limits of interpretation. Inclusion and exclusion errors, programme participation differences, community-ranking disagreement, and geographic allocation measures were treated as distinct outcomes rather than interchangeable indicators of targeting performance. This procedure enabled findings to be traced to their supporting evidence and reduced the risk of drawing conclusions beyond the methodological scope of individual studies. The extraction matrix also retained quality-appraisal concerns to distinguish evidence suitable for primary interpretation from findings requiring conditional or contextual treatment.

2.6. Data Synthesis

The extracted evidence was analyzed using a structured thematic narrative synthesis. Rather than summarizing studies individually, the synthesis examined convergent findings, contrasting results, methodological differences, and contextual conditions influencing welfare classification and social assistance allocation. To preserve differences in the nature of the evidence, studies were organized into five primary analytical strands: actual programme targeting, model-based targeting performance, geographic distributive equity, dynamic vulnerability, and governance and implementation. These strands comprised four, three, two, four, and five studies, respectively. Each study was assigned one primary analytical strand for descriptive purposes, although its findings could contribute to multiple research questions.

Cross-study comparisons were conducted with attention to welfare definitions, reference benchmarks, analytical units, programme objectives, and study designs. Results based on actual beneficiary receipt were distinguished from simulated targeting errors, predictive classification performance, and geographically aggregated allocation outcomes. Convergent and divergent findings were interpreted in relation to these methodological characteristics rather than through statistical pooling or simple counts of significant results. Core-evidence studies

provided the principal basis for substantive interpretation, while conditional-evidence studies were incorporated within their documented methodological limitations. Supporting literature informed contextual and conceptual interpretation without being used to establish primary targeting-performance estimates.

The thematic findings were subsequently integrated to address the five research questions and develop a conceptual framework connecting welfare measurement, eligibility determination, allocation implementation, reassessment, and multidimensional performance evaluation. Governance and institutional capacity were considered cross-cutting conditions influencing these processes. The framework was developed through conceptual integration of the reviewed evidence rather than through empirical estimation of causal relationships. Finally, the synthesized findings were interpreted in relation to Indonesia's social assistance system, particularly the use of welfare deciles, beneficiary eligibility, registry updating, and administrative governance within the Data Tunggal Sosial dan Ekonomi Nasional (DTSEN). These implications were formulated as evidence-informed considerations requiring contextual validation rather than direct empirical conclusions about the effectiveness of DTSEN.

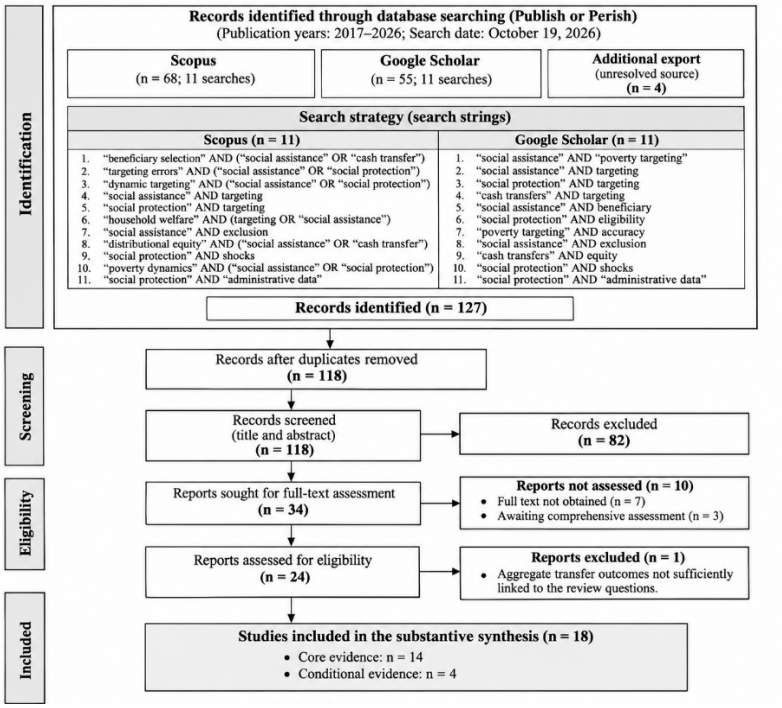

Image: PRISMA 2020 Diagram Flow

RESULTS AND DISCUSSION

Characteristics of the Reviewed Studies

The working synthesis comprised 18 studies retained following full-text assessment, methodological appraisal, and evidence-use adjudication. Of these, 14 were classified as core evidence and four as conditional evidence. Five additional studies were retained for contextual or methodological support but were not used to establish the primary substantive findings. These classifications reflect the methodological suitability of the evidence for particular analytical purposes rather than definitive judgements of overall study quality. The 18-study corpus remains provisional pending the completion of outstanding eligibility assessments and final reconciliation of the study-selection process.

The reviewed studies examined social assistance targeting and allocation across Africa, Asia, and Latin America, encompassing diverse institutional arrangements and socioeconomic conditions. Their methodological approaches included observational investigations, qualitative case studies, mixed-methods analyses, experimental evaluations, longitudinal poverty-transition modelling, and simulation-based targeting assessments. This diversity provided complementary perspectives on the relationship between welfare classification and resource allocation, although substantial differences in welfare indicators, analytical units, and outcome definitions limited direct quantitative comparisons.

The studies were organized into five primary analytical strands: actual programme targeting (n = 4), model-based targeting performance (n = 3), geographic distributive equity (n = 2), dynamic vulnerability (n = 4), and

OPEN ACCESS

governance and implementation (n = 5). Each study was assigned one primary strand for descriptive purposes, while findings relevant to additional research questions were retained during cross-study synthesis.

Table 3. Characteristics of the Reviewed Studies (n = 18)

Author(s) and Year	Country / Setting	Research Design	Targeting Mechanism / Programme
Akyeampong (2024)	Ghana	Qualitative case study	LEAP; geographic targeting, CBT, PMT
Villa & Niño-Zarazúa (2019)	Mexico	Longitudinal poverty-transition modelling	Social protection recertification
Li & Walker (2018)	China	Qualitative case study	Dibao; means-testing and local discretion
Asri (2019)	India	Longitudinal observational analysis	Social pensions; BPL-based eligibility
Lind et al. (2022)	Ethiopia	Mixed methods	PSNP; community and categorical targeting
Follett & Henderson (2023)	Burkina Faso and Indonesia	Predictive model validation	Hybrid PMT–CBT targeting
Abay et al. (2026)	Ethiopia	Cluster-randomized experiment	Cash-transfer allocation; rules and discretion
Eser (2026)	Türkiye	Neighbourhood-level observational analysis	Municipal social assistance allocation
Dadap-Cantal et al. (2021)	Philippines	Qualitative process tracing	Pantawid; Listahanan social registry
Sabates-Wheeler et al. (2022)	Ethiopia	Mixed-methods analysis	PSNP and Humanitarian Food Assistance
Chepngeno-Langat et al. (2023)	Kenya	Observational analysis	Older Persons Cash Transfer
Poirier (2020)	Four Andean countries	Comparative ecological analysis	Geographic allocation of cash transfers
Schnitzer (2019)	Niger	Targeting simulation	PMT, HEA, and geographic targeting
Syafuddin et al. (2026)	Indonesia	Qualitative comparative case study	Social assistance and climate vulnerability
Ray & Oledzka Nielsen (2026)	India	Documentary policy analysis	Shock-responsive social protection
O'Donnell (2026)	Kenya	Qualitative process tracing	Social registry and categorical entitlement
Pruce (2023)	Zambia	Qualitative case study	Social Cash Transfer eligibility
Alix-Garcia et al. (2021)	Mexico	Survey-based targeting simulation	Community leader reporting and asset-based targeting

Note. Core and conditional refer to evidence-use categories established during methodological adjudication. They do not indicate that all core studies have low risk of bias. BPL = Below Poverty Line; CBT = community-based targeting; HEA = household economy analysis; PMT = proxy means testing; PSNP = Productive Safety Net Programme.

The distribution of research designs reveals that social assistance targeting has been examined through different forms of evidence. Observational and experimental studies provide information about actual beneficiary selection and programme allocation, whereas simulation studies assess how alternative targeting mechanisms perform under specified welfare benchmarks and modelling assumptions. Qualitative investigations contribute explanations of administrative practices, local discretion, and institutional constraints but generally do not estimate population-level targeting errors. Consequently, the synthesis emphasizes substantive patterns and methodological differences rather than assuming that all reported targeting outcomes represent comparable measures of programme effectiveness.

OPEN ACCESS

Welfare Classification and Beneficiary Eligibility (RQ1)

The reviewed evidence indicates that welfare classification and beneficiary eligibility involve several related but analytically distinct processes. These include defining the relevant dimensions of socioeconomic deprivation, measuring household circumstances, ranking potential beneficiaries, and applying programme-specific eligibility criteria. Although these processes are often treated as components of a single targeting mechanism, their operational differences influence which households are ultimately considered eligible for assistance.

Studies examining statistical targeting approaches demonstrate the importance of measurement choices in welfare classification. Asri (2019) identified discrepancies between administrative eligibility for social pensions in India and household welfare measured through an asset-based index. Similarly, Schnitzer (2019) demonstrated that targeting methods produced different results when persistent consumption poverty and seasonal food insecurity were used as alternative welfare benchmarks. These findings indicate that a classification method cannot be evaluated independently of the particular form of deprivation it is intended to identify.

The potential value of integrating alternative information sources was illustrated by Follett and Henderson (2023), who examined a hybrid targeting approach combining household characteristics with community-generated welfare rankings. Their findings indicate that statistical prediction can be aligned with locally derived assessments under specified conditions. However, the relevant benchmark was community welfare ranking rather than independently verified consumption poverty, highlighting the importance of distinguishing agreement with local preferences from conventional poverty-targeting accuracy.

Evidence from institutional and qualitative studies further indicates that formal welfare classifications do not independently determine beneficiary eligibility. In Ghana, Akyeampong (2024) documented deviations between prescribed LEAP selection procedures and actual beneficiary nomination practices, particularly through local discretion and political influence. Li and Walker (2018) similarly found that village-level implementation of China's Dibao programme reflected institutional priorities and social relationships extending beyond formal poverty criteria.

Pruce (2023) identified tensions between official eligibility criteria and community perceptions of deservingness in Zambia, particularly regarding the eligibility of households considered capable of working. O'Donnell (2026) demonstrated a different institutional tension in Kenya, where poverty-based registry classification coexisted with categorical entitlement arrangements. The transition towards age-based eligibility for the Older Persons Cash Transfer illustrates how programme objectives can modify the role of household poverty rankings in determining access.

Across these studies, welfare ranking emerges as an important informational component of beneficiary identification rather than a sufficient determinant of eligibility. Actual selection depends on the relationship between measured deprivation, programme objectives, administrative requirements, and institutional practices.

Accuracy, Distributive Equity, and Allocation Efficiency (RQ2)

Evidence from Actual Programme Allocation

Studies examining actual social assistance participation revealed substantial variation in the correspondence between programme receipt and household welfare conditions.

Asri (2019) reported that exclusion error in India's social pension programmes declined from 91.7% in 2004–2005 to 74.3% in 2011–2012. Inclusion error also declined from 57.1% to 40.8%. Although these changes indicate an improvement relative to the study's welfare benchmark, substantial exclusion remained. Importantly, these estimates were calculated using the bottom 40% of an asset-based welfare distribution and should not be interpreted as errors measured against an official national poverty threshold.

In Ethiopia, Lind et al. (2022) identified considerable differences in PSNP participation across livestock-wealth categories. In Afar, 47% of households in the poorest livestock-wealth quintile did not participate in the programme, while 62% of households in the wealthiest quintile participated. These findings reveal a mismatch between programme participation and a particular measure of economic welfare. However, livestock ownership does not capture all forms of vulnerability, and the reported participation differences are not equivalent to independently verified inclusion and exclusion error rates.

Experimental evidence provides a different perspective on allocation decisions. Abay et al. (2026) found that reducing the available transfer budget by approximately 50% decreased the proportion of recipients by around 8–9 percentage points, while reducing the average transfer amount by approximately 43%. The experiment also indicated that greater local discretion increased access among poorer households by approximately 9–11 percentage points relative to a comparable rule-based allocation arrangement. These findings concern an experimental, one-off cash-transfer setting and should not be generalized automatically to regular national social assistance programmes.

OPEN ACCESS

Observational evidence from Kenya showed that household economic characteristics were associated with assistance receipt. Chepngeno-Langat et al. (2023) found that higher household expenditure was negatively associated with receiving the Older Persons Cash Transfer (OR = 0.76), as was engagement in income-generating work (OR = 0.48). These associations suggest that selection was related to observable socioeconomic conditions, but they do not establish the accuracy of the underlying welfare classification or the causal effects of targeting procedures.

Comparative Performance of Targeting Methods

The simulation and model-validation studies demonstrated that the relative performance of targeting methods depends substantially on the welfare benchmark used for evaluation.

The clearest example was provided by Schnitzer (2019). When persistent consumption poverty was used as the benchmark, simulated inclusion error was lower for PMT than for HEA, at 43% and 72%, respectively. However, when targeting was evaluated against lean-season food insecurity, the relationship reversed: HEA produced an inclusion error of 52%, compared with 70% for PMT.

This reversal demonstrates that targeting performance is conditional on the dimension of welfare being measured. A method that performs relatively well in identifying persistent poverty may not be equally effective in identifying seasonal deprivation.

Follett and Henderson (2023) reported out-of-sample disagreement rates of approximately 41% in Burkina Faso and 26% in Indonesia for a hybrid approach predicting community welfare rankings. These estimates demonstrate the predictive performance of the method relative to community preferences rather than its direct accuracy in identifying consumption poverty.

Alix-Garcia et al. (2021) examined the use of community leader information as an alternative to household survey data. Their simulation suggested approximately 77% agreement between classifications, with inclusion and exclusion errors of approximately 17% and 27%, respectively. The study also estimated potential data collection cost savings of approximately 50%. However, these results were based on hypothetical targeting decisions and specified cost assumptions rather than an implemented cash-transfer programme.

Table 4. Comparative Evidence on Targeting Performance and Allocation Outcomes

Study	Evidence Type / Welfare Benchmark	Main Finding	Interpretation Boundary
Asri (2019)	Actual programme receipt; asset-index poverty	Exclusion error declined from 91.7% to 74.3%; inclusion error from 57.1% to 40.8%.	Errors relative to the study's asset-index benchmark.
Lind et al. (2022)	Actual PSNP participation; livestock-wealth quintiles	47% of poorest households did not participate; 62% of wealthiest participated in Afar.	Participation patterns, not formal targeting-error estimates.
Abay et al. (2026)	Experimental allocation	Local discretion increased poorer-household access by approximately 9–11 percentage points.	Context-specific experimental effect.
Chepngeno-Langat et al. (2023)	Observational analysis of benefit receipt	Household expenditure (OR = 0.76) and income-generating work (OR = 0.48) were negatively associated with receipt.	Associations, not classification errors.
Schnitzer (2019)	Simulated targeting; persistent poverty	PMT inclusion error = 43%; HEA = 72%.	Benchmark: persistent consumption poverty.
Schnitzer (2019)	Simulated targeting; seasonal food insecurity	PMT inclusion error = 70%; HEA = 52%.	Benchmark: lean-season food insecurity.
Follett & Henderson (2023)	Hybrid model; community welfare rankings	Disagreement approximately 41% in Burkina Faso and 26% in Indonesia.	Not equivalent to consumption-based poverty errors.
Alix-Garcia et al. (2021)	Survey-based targeting simulation	Approximately 77% agreement; 17% inclusion error; 27% exclusion error.	Hypothetical allocation using asset-based reference information.

Study	Evidence Type / Welfare Benchmark	Main Finding	Interpretation Boundary
Poirier (2020)	Geographic allocation counterfactual	Potential reductions in geographic mistargeting under alternative allocation scenarios.	Simulated allocation, not observed programme improvement.
Eser (2026)	Neighbourhood-level observational analysis	Competitive neighbourhoods received approximately 12% more assistance in an adjacent-area comparison.	Geographic association, not household-level exclusion or causal political favouritism.

Note. OR = odds ratio; PMT = proxy means testing; HEA = household economy analysis. The studies employ different benchmarks, populations, and outcome definitions. Reported estimates should not be directly pooled or used to rank methods across incompatible evaluation criteria.

Geographic Distributive Equity

Geographic analyses revealed an additional dimension of targeting performance that is not captured by household classification accuracy.

Poirier (2020) examined conditional cash transfer allocation across subnational areas in four Andean countries and found that assistance distribution was not consistently aligned with absolute health-related needs. Counterfactual allocation scenarios indicated that alternative approaches could reduce geographic mistargeting. These estimates represent hypothetical improvements under specified allocation rules rather than outcomes of implemented reforms.

Eser (2026) identified an inverted-U relationship between electoral competition and municipal social assistance intensity across neighbourhoods in Istanbul. In one adjacent-area comparison, politically competitive neighbourhoods received approximately 12% more assistance than their counterparts. Although the findings indicate an association between political characteristics and resource distribution, they do not demonstrate household-level targeting errors or establish a causal effect of electoral incentives.

Taken together, these studies suggest that distributive equity requires evaluation at multiple administrative and geographical levels. A welfare classification may be statistically coherent while the resulting allocation of resources remains uneven across populations or territories with different levels of need.

Dynamic Vulnerability and Targeting Responsiveness (RQ3)

The reviewed studies indicate that the relationship between welfare classification and assistance needs changes over time. Household poverty status, livelihood conditions, and exposure to economic or environmental shocks may evolve after the initial collection of socioeconomic information, potentially reducing the relevance of previously established eligibility classifications.

Villa and Niño-Zarazúa (2019) examined poverty transitions and their implications for beneficiary recertification in Mexico. Their model-based analysis suggested recertification intervals of approximately 3.7 years for urban households and 5.1 years for rural households. These estimates reflect model-derived recommendations rather than observed effects of implementing the proposed intervals.

Administrative evidence from the Philippines illustrates a related challenge. Dadap-Cantal et al. (2021) found that beneficiary selection and replacement continued to rely on information from Listahanan 1, collected in 2009, until at least 2019. The prolonged use of older registry information indicates how institutional and administrative constraints can limit the capacity of targeting systems to reflect changing socioeconomic conditions.

A further distinction concerns the relationship between chronic poverty and temporary vulnerability. Sabates-Wheeler et al. (2022) compared Ethiopia's Productive Safety Net Programme with Humanitarian Food Assistance and found that drought exposure was associated with participation in humanitarian assistance (adjusted relative risk ratio = 2.058). The programmes addressed different dimensions of vulnerability, indicating that their allocation patterns should not be assessed against a single undifferentiated poverty benchmark.

Evidence from Indonesia provided additional insight into vulnerability associated with environmental conditions. Syafuddin et al. (2026) found that conventional economic eligibility arrangements did not consistently account for climate exposure and livelihood disruption in the communities examined. However, the qualitative design and reporting limitations prevent these findings from being translated into a nationally representative targeting-error estimate.

Ray and Oledzka Nielsen (2026) documented several approaches to shock-responsive social protection in Odisha, India, including vertical expansion, horizontal expansion, piggybacking on existing programmes, and

OPEN ACCESS

programme refocusing. Existing administrative systems supported rapid delivery to registered beneficiaries but could create difficulties in reaching previously unregistered migrants and informal workers.

Collectively, these findings indicate that the temporal responsiveness of targeting systems constitutes a distinct dimension of allocation performance. The appropriateness of an initial welfare classification depends not only on its measurement accuracy but also on whether eligibility arrangements can respond to subsequent changes in household circumstances.

Integrated Framework of Welfare Targeting and Social Assistance Allocation (RQ4)

The cross-study synthesis informed the development of an integrated conceptual framework linking socioeconomic measurement to actual assistance allocation and subsequent reassessment.

The framework consists of six interconnected analytical stages: defining welfare needs, measuring and ranking household welfare, determining programme eligibility, implementing beneficiary selection and resource allocation, reassessing changing vulnerability, and evaluating targeting and allocation outcomes.

The first stage concerns the definition of the relevant form of deprivation, including persistent poverty, temporary economic insecurity, categorical disadvantage, and exposure to shocks. The second stage translates these definitions into measurable indicators through approaches such as PMT, CBT, HEA, administrative data, or hybrid classification methods.

The third stage involves translating welfare information into programme-specific eligibility rules. These rules may incorporate thresholds, categorical requirements, budget constraints, and administrative documentation.

The fourth stage concerns the implementation of allocation decisions. Findings from Ghana, China, Zambia, and Kenya indicate that institutional practices, community norms, and administrative discretion can influence the relationship between formal classifications and actual beneficiary selection.

The fifth stage addresses reassessment and responsiveness to changing household circumstances. Evidence concerning poverty transitions, registry updating, and shock-responsive assistance demonstrates the relevance of incorporating a temporal dimension into targeting arrangements.

The sixth stage evaluates programme performance using distinct indicators of classification accuracy, distributive equity, administrative efficiency, timeliness, and legitimacy.

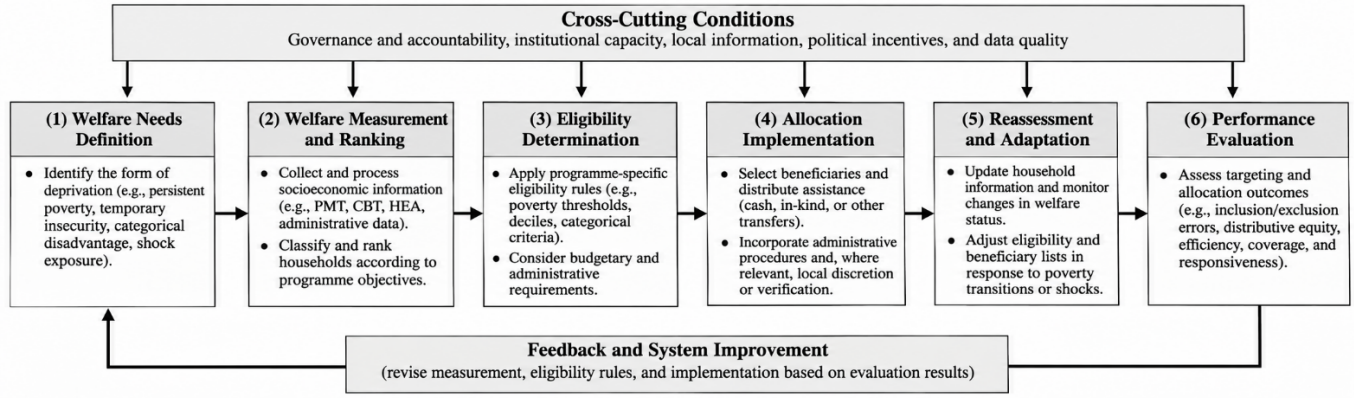

Figure 1. Integrated Framework of Welfare Targeting and Social Assistance Allocation.

Image: Integrated Framework of Welfare Targeting and Social Assistance Allocation

A feedback mechanism connects performance evaluation with subsequent revisions to welfare measurement, eligibility rules, and administrative procedures. Governance, institutional capacity, local information, and political incentives are conceptualized as cross-cutting conditions influencing these processes.

The framework represents an interpretive synthesis of the reviewed literature rather than an empirically validated causal model. Its relationships therefore provide an analytical basis for future investigation rather than established causal effects across social assistance systems.

Governance and Implementation Practices (RQ5)

Governance-related evidence consistently indicated that the implementation of social assistance targeting is influenced by institutional arrangements extending beyond the statistical identification of disadvantaged households.

Akyeampong (2024) documented how political patronage and informal authority influenced beneficiary selection under Ghana's LEAP programme. Similarly, Li and Walker (2018) found that village-level implementation of Dibao in China reflected administrative and social considerations that extended beyond formal poverty criteria.

These findings suggest that discretionary decision-making can create discrepancies between the intended operation of targeting rules and actual beneficiary selection.

However, the experimental findings of Abay et al. (2026) provide an important counterpoint. In the Ethiopian setting examined, greater local discretion improved access among poorer households relative to a comparable rule-based arrangement. This result indicates that local actors may possess relevant information about household disadvantage that standardized eligibility procedures do not fully capture.

Pruce (2023) highlighted another governance mechanism involving the perceived legitimacy of eligibility rules. In Zambia, community judgements concerning deservingness differed from formal programme criteria, particularly regarding the eligibility of households considered capable of working.

O'Donnell (2026) documented institutional tensions surrounding the relationship between poverty-based social registries and categorical entitlement arrangements in Kenya. Dadap-Cantal et al. (2021) further illustrated how administrative constraints and institutional processes affected the updating and use of beneficiary information in the Philippines.

Across these studies, three recurrent governance issues emerged: discrepancies between formal rules and actual selection practices, tensions between standardized classifications and local knowledge, and administrative limitations affecting beneficiary registration and reassessment.

The available evidence suggests that governance arrangements influence the translation of welfare rankings into allocation decisions. Nevertheless, much of this evidence is qualitative or context-specific, and the causal effectiveness of particular accountability, verification, or grievance mechanisms remains insufficiently established.

Cross-Study Synthesis of Key Findings

The synthesis revealed convergence across the reviewed studies regarding the conditional nature of targeting accuracy, the distinction between welfare classification and actual allocation, and the importance of responding to changes in household vulnerability. However, meaningful divergences emerged in the comparative performance of targeting approaches and the consequences of administrative discretion.

These differences were interpreted in relation to the welfare constructs, programme objectives, institutional settings, and methodological designs underlying the reported findings.

Table 3. Cross-Study Synthesis of Key Findings Across RQ1–RQ5

Research Question	Convergent Findings	Divergent Findings / Contradictions	Supporting Studies	Synthesis Conclusion
RQ1. Welfare Classification and Eligibility	Beneficiary identification involves defining need, measuring deprivation, ranking households, and applying eligibility rules.	Formal classifications may conflict with community perceptions of deservingness, local selection practices, or categorical entitlement arrangements.	Asri (2019); Schnitzer (2019); Follett & Henderson (2023); Akyeampong (2024); Li & Walker (2018); Pruce (2023); O'Donnell (2026).	Welfare rankings inform eligibility but do not independently determine beneficiary selection.
RQ2. Targeting Accuracy, Equity, and Efficiency	Performance depends on welfare benchmarks, allocation rules, resource constraints, and units of analysis.	PMT performs better for persistent poverty, whereas HEA performs better for seasonal food insecurity in the Niger simulation. Local discretion improves access in one experiment but is associated with selection distortions in qualitative cases.	Asri (2019); Lind et al. (2022); Schnitzer (2019); Follett & Henderson (2023); Abay et al. (2026); Chepngeno-Langat et al. (2023); Poirier (2020); Eser (2026); Alix-Garcia et al. (2021).	No targeting method demonstrates universal superiority across welfare objectives and allocation contexts.

Research Question	Convergent Findings	Divergent Findings / Contradictions	Supporting Studies	Synthesis Conclusion
RQ3. Dynamic Vulnerability	Poverty transitions, environmental shocks, and registry delays can weaken the correspondence between recorded welfare and current needs.	Approaches differ across periodic recertification, humanitarian assistance, and shock-responsive expansion.	Villa & Niño-Zarazúa (2019); Dadap-Cantal et al. (2021); Sabates-Wheeler et al. (2022); Syafuddin et al. (2026); Ray & Oledzka Nielsen (2026).	Targeting performance should account for changes in household circumstances and administrative responsiveness.
RQ4. Integrated Conceptual Framework	Evidence can be organized into interconnected stages of welfare measurement, eligibility, implementation, reassessment, and evaluation.	The relative importance of these stages varies across programme objectives and institutional settings.	Authors' synthesis informed by Asri (2019); Schnitzer (2019); Villa & Niño-Zarazúa (2019); Abay et al. (2026); Pruce (2023); O'Donnell (2026); Dadap-Cantal et al. (2021), and the wider reviewed corpus.	The proposed framework provides an analytical structure for studying welfare targeting as a multidimensional administrative process.
RQ5. Governance and Implementation	Administrative capacity, local discretion, social norms, and institutional incentives influence allocation decisions.	Discretion may create opportunities for preferential selection but can improve access among poorer households under particular conditions.	Akyeampong (2024); Li & Walker (2018); Pruce (2023); Abay et al. (2026); O'Donnell (2026); Dadap-Cantal et al. (2021); Eser (2026); Ray & Oledzka Nielsen (2026).	Evaluations should consider the relationship between formal eligibility rules, institutional practices, and actual assistance allocation.

Note. Supporting studies contribute evidence to different components of each research question and do not necessarily support every statement within a row. Findings were synthesized narratively because differences in welfare benchmarks, research designs, and outcome measures precluded meaningful statistical pooling. RQ4 represents an author-developed conceptual integration rather than an empirically validated model.

Overall, the synthesis identified three interconnected findings. First, targeting accuracy is conditional on the definition and measurement of welfare. Second, the statistical quality of a classification does not necessarily determine distributive outcomes because eligibility rules and administrative processes intervene between ranking and allocation. Third, household welfare changes over time, potentially reducing the continued relevance of previously established classifications.

These findings provide the basis for interpreting welfare targeting as a multidimensional process involving measurement, institutional decision-making, allocation, and reassessment.

Discussion

Reconsidering Targeting Accuracy as a Conditional Measure of Welfare Identification

The findings challenge the interpretation of targeting accuracy as a universally comparable indicator of social assistance performance. Although inclusion and exclusion errors provide useful information about beneficiary identification, their substantive meaning depends on how welfare deprivation is defined, which populations are considered eligible, and the reference measure against which targeting decisions are evaluated. The reviewed evidence indicates that differences in these methodological choices can substantially alter conclusions about the performance of alternative targeting arrangements.

OPEN ACCESS

The comparison between PMT and HEA in Niger provides a particularly important illustration. Schnitzer (2019) found that PMT produced lower simulated inclusion errors when persistent consumption poverty was used as the reference, whereas HEA performed better when targeting was evaluated against seasonal food insecurity. This reversal indicates that apparent methodological superiority may arise from alignment between the classification approach and the welfare dimension being measured rather than from an inherent advantage of the method itself.

This interpretation is reinforced by studies using alternative reference measures. Follett and Henderson (2023) evaluated a hybrid targeting approach against community welfare rankings, while Alix-Garcia et al. (2021) assessed simulated allocation using household-reported asset information. Although both investigations contribute to understanding the potential of alternative information sources, their reported accuracy measures cannot be interpreted as directly comparable estimates of consumption-based poverty targeting.

The theoretical implication is that targeting accuracy should be understood as a benchmark-dependent property of a classification system rather than an absolute indicator of welfare identification quality. Measures of performance become meaningful only when the intended welfare construct, programme objective, reference population, and consequences of classification errors are explicitly established.

This perspective also suggests that conventional comparisons between targeting methods may be misleading when studies employ different welfare constructs. A method designed to identify persistent economic deprivation should not necessarily be expected to perform equally well in predicting temporary food insecurity, sudden income losses, or socially recognized vulnerability.

Accordingly, the contribution of the review lies partly in clarifying the conditions under which reported targeting outcomes can be compared. Rather than identifying a universally superior method, the synthesis indicates that method selection should be evaluated in relation to the specific forms of disadvantage that social assistance programmes are intended to address.

From Welfare Rankings to Allocation Outcomes: Explaining the Classification–Allocation Gap

A second important finding concerns the distinction between classifying household welfare and allocating social assistance. Much of the targeting literature evaluates whether households can be ranked accurately according to estimated socioeconomic conditions. However, the present synthesis indicates that these rankings represent only one component of the broader allocation process.

Evidence from India illustrates this distinction. Asri (2019) identified substantial discrepancies between asset-based welfare status and social pension receipt, indicating that the identification of economically disadvantaged households does not necessarily translate into their inclusion in benefit programmes. Similar concerns emerged in Ethiopia, where Lind et al. (2022) documented differences between PSNP participation and livestock-based welfare rankings.

These discrepancies cannot be attributed exclusively to measurement error. Programme eligibility requirements, administrative documentation, budgetary restrictions, and local implementation practices may also influence which households receive assistance.

The governance studies provide additional explanations for this classification–allocation gap. Akyeampong (2024) and Li and Walker (2018) demonstrated that beneficiary selection may reflect political relationships, local institutional priorities, and discretionary decision-making. O'Donnell (2026) further showed that the role of poverty rankings can be modified by categorical eligibility arrangements, suggesting that different programme objectives may produce different relationships between estimated welfare and entitlement.

Taken together, these findings support a distinction between three analytical dimensions: welfare measurement quality, eligibility determination, and realized benefit allocation.

The first concerns whether socioeconomic information adequately represents the intended form of deprivation. The second concerns how programme rules convert welfare information into eligibility decisions. The third concerns whether those decisions result in assistance reaching the intended recipients.

These dimensions are interconnected but should not be treated as interchangeable. A classification system may perform well against an independent poverty benchmark while assistance allocation remains constrained by eligibility rules, incomplete registration, or institutional implementation. Conversely, observed differences between welfare rankings and benefit receipt may reflect legitimate categorical programme objectives rather than errors in welfare measurement.

This distinction extends the interpretation of targeting performance beyond statistical prediction. It suggests that evaluations should examine the sequence through which welfare information becomes an administrative decision and subsequently an observed allocation outcome.

OPEN ACCESS

The conceptual contribution is therefore not simply that administrative implementation matters, but that classification accuracy and allocation performance constitute distinct evaluation targets requiring different data, benchmarks, and inferential approaches.

Distributive Equity and the Ambivalent Role of Local Discretion

The findings further reveal that social assistance allocation involves trade-offs between standardized decision-making and the use of local information. Standardized welfare classifications offer procedural consistency and potentially reduce arbitrary variation in eligibility assessment. However, they may not fully capture temporary hardship, informal livelihoods, or dimensions of vulnerability that are more readily observable within communities.

The evidence concerning local discretion illustrates this tension. Qualitative studies from Ghana and China associated discretionary practices with deviations from formal eligibility procedures, including political patronage and the use of assistance allocation for local institutional purposes (Akyeampong, 2024; Li & Walker, 2018).

In contrast, Abay et al. (2026) found that greater local discretion increased access among poorer households in an experimental cash-transfer allocation setting. The difference between these findings should not be interpreted as a direct empirical contradiction, because the studies examined different institutional arrangements, allocation rules, and outcome measures.

Nevertheless, the comparison raises an important explanatory question: under what conditions does local discretion improve the identification of disadvantaged households, and when does it create opportunities for preferential selection?

One plausible interpretation is that local discretion performs different functions depending on the institutional setting. Where standardized information is incomplete, community knowledge may provide additional evidence about household needs. Where accountability is weak or decisions are influenced by political relationships, discretionary authority may instead weaken the correspondence between programme objectives and actual beneficiary selection.

This explanation remains a synthesis-based proposition rather than an empirically established mechanism across the reviewed studies. The evidence supports the relevance of institutional context but does not identify a universally applicable combination of administrative safeguards that would ensure beneficial discretionary decisions.

The findings concerning geographic allocation reinforce the importance of considering equity beyond individual welfare rankings. Poirier (2020) demonstrated that allocation across subnational areas may not correspond consistently to absolute health-related needs, while Eser (2026) identified associations between political competition and the geographic distribution of assistance.

These studies suggest that a technically consistent household classification system does not necessarily guarantee equitable allocation across territories. However, geographically unequal distribution is not automatically evidence of unfair targeting, because programme objectives and implementation costs may legitimately differ across locations.

Accordingly, distributive equity should be evaluated relative to explicitly defined policy objectives, population needs, and allocation constraints.

The broader implication is that welfare targeting involves more than minimizing prediction errors. It also requires understanding how institutional rules, budgetary decisions, local information, and geographic priorities influence the distribution of limited resources.

Dynamic Vulnerability and the Temporal Limitations of Welfare Rankings

The third cross-cutting finding concerns the limitations of welfare classifications based on information collected at a single point in time. Conventional targeting arrangements typically rely on socioeconomic characteristics observed during registration or survey enumeration. However, these characteristics may change because of employment disruption, household transitions, migration, or exposure to environmental and economic shocks.

The longitudinal analysis by Villa and Niño-Zarazúa (2019) demonstrates the relevance of poverty transitions for reassessment decisions. Their findings suggest that the appropriate timing of recertification may differ between household populations, although the proposed intervals were derived from modelling rather than evaluated as implemented reforms.

Dadap-Cantal et al. (2021) identified a different form of temporal limitation in the Philippines, where the continued use of older registry information constrained the responsiveness of programme administration to changing circumstances.

These studies suggest that the accuracy of welfare classification cannot be considered independently of the time elapsed between measurement and the allocation decision.

The evidence from Ethiopia further indicates that different forms of vulnerability may require different eligibility arrangements. Sabates-Wheeler et al. (2022) found that drought exposure was associated with humanitarian assistance receipt, illustrating the distinction between programmes addressing persistent food insecurity and those responding to temporary shocks.

Qualitative and documentary studies from Indonesia and India extended this issue to environmental vulnerability and compound crises (Syafuddin et al., 2026; Ray & Oledzka Nielsen, 2026). Although their designs do not establish the causal effectiveness of particular targeting reforms, they demonstrate how established beneficiary lists and administrative rules may become insufficient when vulnerability emerges outside previously registered populations.

These findings suggest that targeting performance possesses a temporal dimension that is not adequately represented by conventional cross-sectional accuracy measures.

A classification can correspond reasonably well to household conditions when originally constructed but become less informative as circumstances change. Conversely, a targeting system may incorporate initially imperfect classifications yet improve its alignment with current needs through effective reassessment and verification.

This distinction has implications for how welfare targeting is conceptualized. Targeting quality should be evaluated not only through initial classification performance but also through the responsiveness of eligibility and allocation procedures to changing household circumstances.

The review does not establish that more frequent updating necessarily improves programme outcomes. Updating introduces administrative costs, potential measurement inconsistencies, and risks of inappropriate changes to beneficiary status. Its value therefore depends on whether new information improves the identification of relevant needs and is translated into timely and appropriate allocation decisions.

These considerations support further investigation of the relationship between welfare mobility, registry updating, programme responsiveness, and allocation outcomes.

Conceptual Contribution: Integrating Measurement, Allocation, and Institutional Responsiveness

The synthesis offers an integrated perspective on welfare targeting by bringing together forms of evidence that are frequently examined separately. Model-validation studies primarily assess the quality of welfare measurement, programme evaluations examine beneficiary selection and benefit receipt, geographic analyses investigate distributive patterns, and governance studies explain institutional implementation.

Although each perspective contributes valuable evidence, none independently captures the complete process through which socioeconomic information is translated into assistance allocation.

The integrated framework developed in this review addresses this conceptual fragmentation by distinguishing six analytical stages: welfare-need definition, measurement and ranking, eligibility determination, allocation implementation, reassessment, and performance evaluation.

The framework identifies several relationships that warrant further examination.

First, the relationship between welfare rankings and programme eligibility is conditional on how deprivation is defined and how programme rules translate rankings into administrative decisions.

Second, the relationship between eligibility and benefit receipt may vary with administrative capacity, local discretion, resource constraints, and institutional accountability.

Third, the relationship between recorded welfare and actual household needs may change over time, making reassessment and adaptive eligibility relevant components of targeting performance.

These relationships provide an analytical structure for understanding why improvements in predictive classification do not necessarily produce corresponding improvements in distributive outcomes.

The framework should not, however, be interpreted as evidence that all six stages form a universally applicable causal sequence. The reviewed studies do not jointly test the entire model, and the relationships between its components have not been estimated using a common empirical design.

Its contribution is therefore primarily conceptual and methodological: it organizes heterogeneous findings into a coherent structure while identifying the distinctions and connections that future empirical research should examine.

For social assistance research, this perspective encourages evaluations that follow the administrative process from the measurement of need to actual allocation and subsequent reassessment, rather than relying exclusively on the predictive performance of household welfare classifications.

Implications for Indonesia and Limitations of the Evidence

The findings are particularly relevant to Indonesia's use of integrated socioeconomic information through the Data Tunggal Sosial dan Ekonomi Nasional (DTSEN). Welfare deciles provide a standardized means of representing relative socioeconomic positions, but their usefulness for social assistance allocation depends on how the rankings correspond to programme objectives, eligibility requirements, and actual household circumstances.

The international evidence suggests that evaluating a national welfare registry solely through data integration or classification consistency would provide an incomplete assessment of targeting performance. A more informative evaluation would distinguish the validity of welfare rankings, the appropriateness of eligibility rules, the accuracy of beneficiary selection, and the responsiveness of administrative procedures to changing needs.

This distinction is important for programmes that combine socioeconomic eligibility with categorical requirements or operate under different resource constraints.

The findings also suggest the relevance of examining how local verification procedures, administrative coordination, and registry updating influence the translation of welfare information into actual benefit receipt. However, these implications represent contextual applications of international evidence rather than direct empirical assessments of DTSEN.

The current evidence base does not establish whether DTSEN has improved or reduced inclusion errors, exclusion errors, distributive equity, or programme efficiency. Such conclusions require direct evaluation using appropriately defined welfare benchmarks, administrative records, and credible research designs.

Several limitations should also be considered when interpreting the synthesis. First, the reviewed studies employed heterogeneous research designs and welfare measures, limiting the comparability of targeting outcomes. Second, evidence from simulations and predictive models cannot establish actual programme effectiveness without appropriate field validation. Third, qualitative governance studies provide contextually rich explanations of implementation practices but generally do not support representative estimates of the prevalence of targeting errors.

Fourth, much of the evidence concerning registry updating and shock-responsive assistance remains observational, model-based, or documentary, restricting conclusions about the causal effects of particular reforms. Finally, the synthesis is based on a provisionally adjudicated corpus of 18 studies, with outstanding eligibility verification and independent appraisal procedures still to be completed.

Despite these limitations, the synthesis provides a coherent explanation of why welfare targeting should not be evaluated solely through classification accuracy. Its performance depends on the relationship between welfare measurement, administrative eligibility, actual allocation, institutional governance, and the capacity to respond to changing household vulnerability.

CONCLUSION

This systematic literature review examined how welfare rankings influence social assistance allocation through targeting accuracy, distributive equity, dynamic vulnerability, and institutional governance. Drawing on a provisional synthesis of 18 studies, the findings indicate that targeting performance cannot be assessed solely through the statistical accuracy of socioeconomic classifications. The effectiveness of targeting methods depends on how household deprivation is defined and measured, while actual benefit allocation is also shaped by programme eligibility, administrative procedures, local discretion, and institutional conditions. The evidence further demonstrates that targeting approaches may perform differently across welfare benchmarks and that classifications based on static household information may become less representative as socioeconomic circumstances change. By integrating these findings, the review proposes a six-stage conceptual framework connecting welfare-need definition, measurement and ranking, eligibility determination, allocation implementation, reassessment, and performance evaluation. This framework advances an understanding of social assistance targeting as a multidimensional process in which classification accuracy, distributive outcomes, and institutional responsiveness must be evaluated together.

For Indonesia, these findings highlight the importance of evaluating the Data Tunggal Sosial dan Ekonomi Nasional (DTSEN) beyond the consistency of welfare deciles and the integration of socioeconomic records. Improving social assistance targeting requires attention to whether welfare rankings correspond to actual deprivation, how eligibility decisions translate into benefit receipt, and whether administrative systems respond to changing household vulnerability. Future research should prioritize independent validation of welfare classifications, longitudinal assessments of welfare mobility, and rigorous evaluations of targeting and verification mechanisms. Nevertheless, differences in research designs, welfare benchmarks, and institutional contexts limit the generalizability of the reviewed findings, while the proposed framework remains subject to empirical validation. The central implication is that effective social assistance allocation depends not merely on identifying disadvantaged

households accurately, but on translating welfare information into equitable, accountable, and responsive assistance decisions.

REFERENCES

Abay, K. A., Berhane, G., Gilligan, D. O., Tafere, K., & Taffesse, A. S. (2026). Targeting social assistance in fragile settings: An experiment in community-based targeting. *Journal of Development Economics, 182*, Article 103819. https://doi.org/10.1016/j.jdeveco.2026.103819

Aiken, E., Bellue, S., Karlan, D., Udry, C., & Blumenstock, J. E. (2022). Machine learning and phone data can improve targeting of humanitarian aid. *Nature, 603*, 864–870. https://doi.org/10.1038/s41586-022-04484-9

Akyeampong, B. (2024). The politics of beneficiary selection: A case study of Ghana's Livelihood Empowerment Against Poverty (LEAP) cash transfer programme. *Global Social Policy, 24*(3), 390–411. https://doi.org/10.1177/14680181241256471

Alix-Garcia, J. M., Sims, K. R. E., & Costica, L. (2021). Better to be indirect? Testing the accuracy and cost-savings of indirect surveys. *World Development, 142*, Article 105419. https://doi.org/10.1016/j.worlddev.2021.105419

Asri, V. (2019). Targeting of social transfers: Are India's poor older people left behind? *World Development, 115*, 46–63. https://doi.org/10.1016/j.worlddev.2018.11.001

Badan Pusat Statistik. (2026, August 22). *DTSEN jadi rujukan bersama, BPS jelaskan arti desil*. https://www.bps.go.id/id/news/2026/08/22/938/dtsen-jadi-rujukan-bersama--bps-jelaskan-arti-desil.html

Banerjee, A., Hanna, R., Olken, B. A., Satriawan, E., & Sumarto, S. (2023). Electronic food vouchers: Evidence from an at-scale experiment in Indonesia. *American Economic Review, 113*(2), 514–547. https://doi.org/10.1257/aer.20210461

Brown, C., Ravallion, M., & van de Walle, D. (2018). A poor means test? Econometric targeting in Africa. *Journal of Development Economics, 134*, 109–124. https://doi.org/10.1016/j.jdeveco.2018.05.004

Chepngeno-Langat, G., van der Wielen, N., Falkingham, J., & Evandrou, M. (2023). Targeting cash transfers on the "poorest of the poor" in the slums: How well did the Kenya's Older Persons Cash Transfer Programme perform? *Journal of Aging & Social Policy, 35*(1), 107–124. https://doi.org/10.1080/08959420.2021.1926197

Dadap-Cantal, E. L., Fischer, A. M., & Ramos, C. G. (2021). Targeting versus social protection in cash transfers in the Philippines: Reassessing a celebrated case of social protection. *Critical Social Policy, 41*(3), 364–384. https://doi.org/10.1177/02610183211009891

Dietrich, S., Malerba, D., & Gassmann, F. (2024). Predicting social assistance beneficiaries: On the social welfare damage of data biases. *Data & Policy, 6*, Article e3. https://doi.org/10.1017/dap.2023.38

Eser, C. F. (2026). Distributive politics in opposition: Electoral targeting of social assistance in Istanbul. *European Journal of Political Economy, 95*, Article 102896. https://doi.org/10.1016/j.ejpoleco.2026.102896

Follett, L., & Henderson, H. (2023). A hybrid approach to targeting social assistance. *Journal of Development Economics, 160*, Article 103002. https://doi.org/10.1016/j.jdeveco.2022.103002

Grosh, M., Leite, P., Wai-Poi, M., & Tesliuc, E. (Eds.). (2022). *Revisiting targeting in social assistance: A new look at old dilemmas*. World Bank.

Hanna, R., & Olken, B. A. (2018). Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives, 32*(4), 201–226. https://doi.org/10.1257/jep.32.4.201

Henderson, H., & Follett, L. (2022). Targeting social safety net programs on human capabilities. *World Development, 151*, Article 105741. https://doi.org/10.1016/j.worlddev.2021.105741

Hillebrecht, M., Klonner, S., & Pacere, N. A. (2023). The dynamics of poverty targeting. *Journal of Development Economics, 161*, Article 103033. https://doi.org/10.1016/j.jdeveco.2022.103033

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285–291. https://doi.org/10.3233/EFI-180221

International Labour Organization. (2024). *World social protection report 2024–26: Universal social protection for climate action and a just transition*.

Kementerian Sosial Republik Indonesia. (2025). *Peraturan Menteri Sosial Nomor 3 Tahun 2025 tentang Pemutakhiran dan Penggunaan Data Tunggal Sosial dan Ekonomi Nasional untuk Bantuan Sosial,*

OPEN ACCESS

Pemberdayaan Sosial, dan Program Penyelenggaraan Kesejahteraan Sosial. https://peraturan.bpk.go.id/Details/323075/permensos-no-3-tahun-2025

Li, M., & Walker, R. (2018). Targeting social assistance: Dibao and institutional alienation in rural China. *Social Policy & Administration, 52*, 771–789. https://doi.org/10.1111/spol.12261

Lind, J., Sabates-Wheeler, R., Hoddinott, J., & Taffesse, A. S. (2022). Targeting social transfers in Ethiopia's agro-pastoralist and pastoralist societies. *Development and Change, 53*(2), 279–307. https://doi.org/10.1111/dech.12694

Mariyah, S., & Wobcke, W. (2025). Selective guided domain adaptation for improving the proxy means test for poverty targeting. *International Journal of Data Science and Analytics, 20*, 5999–6017. https://doi.org/10.1007/s41060-025-00802-2

O'Donnell, K. (2026). Contested visions for social protection in Kenya: The Older Persons Cash Transfer and the social registry. *Social Inclusion, 14*, Article 11219. https://doi.org/10.17645/pag.11219

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Poirier, M. J. P. (2020). Geographic targeting and normative frames: Revisiting the equity of conditional cash transfer program distribution in Bolivia, Colombia, Ecuador, and Peru. *International Journal for Equity in Health, 19*, Article 125. https://doi.org/10.1186/s12939-020-01233-0

Presiden Republik Indonesia. (2025). *Instruksi Presiden Nomor 4 Tahun 2025 tentang Data Tunggal Sosial dan Ekonomi Nasional*. https://peraturan.bpk.go.id/Details/314649/inpres-no-4-tahun-2025-data-tunggal-sosial-dan-ekonomi-nasional

Pruce, K. (2023). The politics of who gets what and why: Learning from the targeting of social cash transfers in Zambia. *The European Journal of Development Research, 35*(4), 820–839. https://doi.org/10.1057/s41287-022-00540-2

Ray, S., & Oledzka Nielsen, M. (2026). Navigating concurrent shocks: A case study of shock-responsive social protection in India. *Journal of Human Rights and Social Work, 11*, 199–211. https://doi.org/10.1007/s41134-026-00438-7

Republik Indonesia. (2022). *Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi*. https://peraturan.bpk.go.id/Details/229798/uu-no-27-tahun-2022

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews, 10*, Article 39. https://doi.org/10.1186/s13643-020-01542-z

Sabates-Wheeler, R., Hirvonen, K., Lind, J., & Hoddinott, J. (2022). Expanding social protection coverage with humanitarian aid: Lessons on targeting and transfer values from Ethiopia. *The Journal of Development Studies, 58*(10), 1981–2000. https://doi.org/10.1080/00220388.2022.2096443

Schnitzer, P. (2019). How to target households in adaptive social protection systems? Evidence from humanitarian and development approaches in Niger. *The Journal of Development Studies, 55*(sup1), 75–90. https://doi.org/10.1080/00220388.2019.1687877

Schnitzer, P., & Stoeffler, Q. (2024). Targeting social safety nets: Evidence from nine programs in the Sahel. *The Journal of Development Studies, 60*(4), 574–595. https://doi.org/10.1080/00220388.2023.2291325

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., . . . Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ, 366*, Article l4898. https://doi.org/10.1136/bmj.l4898

Syafuddin, K., Irmayani, N. R., Habibullah, H., Setiawan, H. H., Susantyo, B., Sumarno, S., Murni, R., Purwasantana, D., Ganti, M., Arifin, J., Aldyan, R. A., Nathanael, G. K., Hapsery, A., & Yuda, T. K. (2026). Adaptive social protection and risk communication strategies for addressing disaster and climate vulnerability in rural Indonesia. *Discover Sustainability, 7*, Article 1411. https://doi.org/10.1007/s43621-026-03606-x

Villa, J. M., & Niño-Zarazúa, M. (2019). Poverty dynamics and graduation from conditional cash transfers: A transition model for Mexico's Progresa-Oportunidades-Prospera program. *The Journal of Economic Inequality, 17*, 219–251. https://doi.org/10.1007/s10888-018-9399-5

Wobcke, W., & Mariyah, S. (2023). Machine learning and data augmentation in the proxy means test for poverty targeting. *Statistical Journal of the IAOS, 39*(4), 961–977. https://doi.org/10.3233/SJI-230033
World Bank. (2025). *State of social protection report 2025: The 2-billion-person challenge*.