

RESEARCH

Open Access



Development of indicators to evaluate the effectiveness of antimicrobial stewardship programs in Korea: a Delphi consensus study

Sung Min Lim^{1†}, Raeseok Lee^{2†}, Jinnam Kim^{3,4}, Ji Young Park⁵, Choseok Yoon³, Myung Jin Lee⁶, Song Mi Moon⁷, Yong Chan Kim⁸, Hong Bin Kim⁷, Se Yoon Park^{3,4*} and Bongyoung Kim^{3,4*}

Abstract

Background A pilot program launched in November 2024 to establish and operationalize Antimicrobial Stewardship Programs (ASPs) in Korean hospitals is expected to accelerate nationwide adoption. However, consensus on indicators for evaluating ASP performance remains lacking. This study aimed to develop indicators for evaluating the short- and long-term effectiveness of ASPs.

Methods An expert panel comprised 18 specialists from academic societies and government agencies involved in infectious diseases and antimicrobial control in South Korea. Candidate indicators for evaluating ASP effectiveness were identified by five infectious disease specialists through a systematic literature review. The Delphi survey consisted of an initial open-ended round, followed by two multiple-choice rounds (rounds 2 and 3) with an intervening online panel meeting. Responses to the multiple-choice questions were rated on a 7-point Likert scale, and indicators with a content validity ratio of ≥ 0.47 were selected.

Results The systematic literature review yielded 59 candidate indicators of ASP effectiveness, which were classified into seven primary categories. After the third-round Delphi survey, the final set of indicators included the following. In the Antibiotic Use Indicators category: (1) days of therapy per 1,000 patient-days, (2) defined daily dose per 1,000 patient-days, (3) reduction rate of broad-spectrum antibiotic use, (4) intravenous-to-oral antibiotic conversion rate, (5) reduction rate of unnecessary combination therapy, and (6) change in antibiotic prescription ratio according to the World Health Organization Access, Watch, and Reserve classification. In the Process-Related Indicators category: (1) appropriate rate of preoperative surgical antibiotic prophylaxis, (2) implementation of therapeutic drug monitoring with dosage recommendations, and (3) establishment of institutional antibiotic use guidelines. In the Department-Specific Indicators related to operating rooms and surgery: (1) appropriate rate of preoperative surgical antibiotic prophylaxis, (2) appropriateness of postoperative antibiotic discontinuation timing, and (3) appropriateness of

[†]Sung Min Lim and Raeseok Lee have contributed equally as first authors.

*Correspondence:
Se Yoon Park
syParkmd@hanyang.ac.kr
Bongyoung Kim
sobakas@hanyang.ac.kr

Full list of author information is available at the end of the article



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

prophylactic antibiotic selection. Finally, among the Other Indicators, the implementation rate of core ASP elements was selected as an indicator for evaluating ASP effectiveness.

Conclusions This study established consensus-based indicators and proposed short- and long-term targets for evaluating the effectiveness of ASPs in Korea. These indicators enable standardized assessment and help inform evidence-based policies for the sustainable operation of ASPs.

Keywords Antimicrobial stewardship, Indicators, Republic of Korea, Evaluation, Delphi method

Background

Antimicrobial resistance poses a growing global health crisis, driving substantial morbidity, mortality, and healthcare costs [1, 2]. Antimicrobial stewardship programs (ASPs) have been widely endorsed by major health organizations as an essential strategy to optimize antibiotic use, improve patient outcomes, and mitigate the emergence of resistance [3, 4]. To ensure that ASP efforts are effective and sustainable, they must be supported by robust performance metrics that demonstrate impact, guide continuous improvement, and enable benchmarking across institutions.

Several countries have adopted indicators to evaluate ASP effectiveness, including days of therapy (DOT), defined daily dose (DDD), incidence of *Clostridioides difficile* infection (CDI), and adherence to prescription guidelines [5–7]. However, these measures are insufficient for comprehensive evaluation, as they capture only a subset of ASP activities. Recent international efforts have highlighted the importance of structured consensus building to identify relevant and practical indicators, with methods such as modified Delphi surveys used to reach agreement among multidisciplinary experts [8, 9].

In Korea, antibiotic consumption remains among the highest compared with other high-income countries, while antimicrobial resistance continues to rise [10]. The Korean National Action Plan on antimicrobial resistance has prioritised ASP at the policy level [11]; however, implementation at the hospital level has been heterogeneous. Although previous studies have reported favorable impacts on antibiotic consumption, resistance trends [12–14] and have provided valuable insights, they do not establish a harmonized framework for evaluating ASP performance at the national level. In November 2024, the Korean government launched a reimbursement-based pilot program involving 78 hospitals, representing a major step toward the formal institutionalization of ASPs [15]. Early implementation revealed substantial heterogeneity in infrastructure, staffing, and outcome measurements across institutions, reinforcing the need for a unified evaluation framework. As the program transitions toward broader implementation, a national consensus on evaluation indicators has become increasingly important to ensure fair performance assessment and sustainable reimbursement mechanisms. In the absence

of standardized metrics, reliable national benchmarking and long-term policy evaluation remain challenging.

In light of these considerations, we conducted a structured Delphi consensus process to develop a set of performance indicators tailored to the Korean healthcare system and potentially adaptable to other resource-limited settings. Drawing on published evidence and expert consensus from clinicians, pharmacists, microbiologists, and policymakers—including representatives from national public health and reimbursement agencies (the Korea Disease Control and Prevention Agency and the Health Insurance Review and Assessment Service)—we propose metrics that are practical within existing data systems and aligned with international standards. These indicators aim to provide a coherent framework for monitoring hospital-level ASPs, evaluating national initiatives, and informing ASP implementation in comparable healthcare contexts.

Methods

Overview and assembly of panel

We conducted a systematic literature review to identify potential key quality indicators (QIs), three sequential rounds of online surveys, and one online expert panel meeting. Figure 1 presents an overview of the modified Delphi method used in this study. The consensus process synthesized the independent judgments of a multidisciplinary expert panel. A total of 18 panelists participated, representing the Korean Society of Infectious Diseases ($n=4$), the Korean Society for Antimicrobial Therapy ($n=4$), the Korean Society of Pediatric Infectious Diseases ($n=2$), the Korea National Antimicrobial Use Analysis System ($n=2$), the Korean Society of Health-System Pharmacists ($n=2$), the Korea Disease Control and Prevention Agency ($n=2$), and the Health Insurance Review and Assessment Service ($n=2$).

Search strategy

A systematic literature search was performed to identify and analyze potential QIs used to evaluate the effectiveness of ASPs in countries with established stewardship systems, including the United States, Japan, the United Kingdom, and Australia. These countries were selected because they have formally institutionalized ASPs and have accumulated a relatively large body of research on

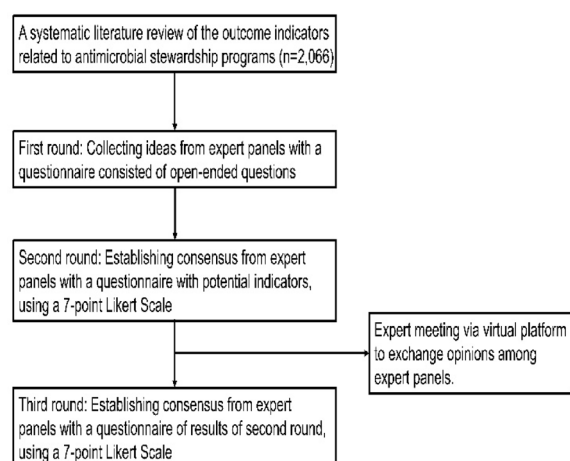


Fig. 1 Flowchart of the study procedures

ASP evaluation indicators, making their frameworks particularly relevant as references for developing a Korean ASP evaluation system. The review focused on methods for monitoring ASP implementation and assessing its short- or long-term effectiveness.

Studies written in English that addressed the development or evaluation of ASP-related indicators were included. The following were excluded: (1) studies not involving humans; (2) studies primarily focused on human immunodeficiency virus infection; and (3) editorials, comments, case reports, errata, and conference abstracts (Additional file 1).

The literature search was conducted on November 22, 2024, and the identified records were reviewed in full text by four infectious disease specialists and one pediatric infectious disease specialist (C.Y., J.K., M.J.L., R.L., and J.Y.P.) between October 28 and December 29, 2024. Any discrepancies regarding eligibility were resolved through discussion and consensus among the reviewers. Following full-text review, all potential QIs proposed or applied in the literature were extracted to develop a preliminary list of candidate indicators (Additional file 2). Each QI was categorized into one of seven domains: antibiotic use, clinical outcome, microbiological, economic, process-related, department-specific, or other indicators. Process-related indicators were defined as measures reflecting hospital-wide ASP monitoring activities, capturing the overall adequacy of antibiotic management practices across the institution. Department-specific indicators, in contrast, were defined as measures targeting granular, component-level evaluation of antibiotic use within high-priority clinical settings, such as the operating room and surgical ward, and were designed to provide more operationally targeted feedback at the point of care. Although certain department-specific indicators share a common focus with process-related indicators, they were retained as distinct categories to enable

complementary evaluation at different levels of institutional practice. Duplicate items were removed and conceptually overlapping indicators were consolidated into unified descriptions.

Development of survey items and the process of Delphi method

Each survey was distributed to the expert panel via email, and the preliminary candidate QIs derived from the literature review were shared with the panelists in advance. The first-round survey (February 12–24, 2025) used an open-ended questionnaire, allowing panel members to suggest revisions or additions to indicators and provide free-text comments. The second-round survey (March 11–19, 2025) consisted of a structured, closed-ended questionnaire containing the same items as those in the first round. Panelists evaluated each indicator according to three criteria: (1) expected outcomes reflecting appropriate antibiotic use, (2) the degree of improvement anticipated when ASP activities are led by infectious disease specialists compared with other clinicians, and (3) feasibility of assessment within domestic healthcare institutions (Additional file 3).

After the second-round survey, an online expert meeting was held on March 21, 2025, to review the results and discuss items requiring further clarification or consensus. The third-round survey (April 1–7, 2025) incorporated revisions based on the preceding discussion and aimed to finalize indicator selection. Evaluation criteria for this round included: (1) expected outcomes reflecting appropriate antibiotic use; (2) feasibility of measurement in domestic institutions; and (3) the need for generating evidence to support ASP implementation. Items in the second- and third-round surveys were rated on a 7-point Likert scale ranging from 1 (“strongly disagree”) to 7 (“strongly agree”), with scores of 6 or 7 defined as essential responses. In the third-round survey, an additional question item regarding “an appropriate timeframe for achieving measurable outcomes” was included, requiring respondents to select one of four options: short-term (3–4 years), mid-term (5–9 years), long-term (≥ 10 years), or not applicable.

Validation of items

Responses from the second and third Delphi survey rounds were analyzed using the content validity ratio (CVR). The CVR was calculated using the formula $CVR = (N_e - N/2) / (N/2)$, where N_e represents the number of experts who rated an item as “essential,” and N denotes the total number of experts [16]. Items with a $CVR \geq 0.47$ were accepted, whereas those with a $CVR < 0.40$ were further reviewed through panel discussion. The threshold of $CVR \geq 0.47$ was based on the corrected critical value for a panel of 18 experts at a significance level of $p < 0.05$,

as derived by Ayre and Scally [17], who identified a computational error in Lawshe's original method and provided revised critical values. The secondary threshold of $\text{CVR} \geq 0.40$ corresponds to a significance level of $p < 0.10$ and was applied to flag borderline items for structured panel discussion, rather than automatic rejection [16, 17]. Alternatively, items with a lower limit of the 95% confidence interval of ≥ 6.5 were accepted, and those with a lower limit of ≥ 5.5 were considered for discussion. The two thresholds were established a priori through deliberation among the research team before data collection, as no established reference exists for these specific cutoffs. A lower bound of ≥ 6.5 indicates that the conservative estimate of the confidence interval falls within the "essential" range of the 7-point Likert scale (score ≥ 6), reflecting strong consensus with minimal inter-rater variability, and was therefore adopted as the acceptance criterion. A lower bound of ≥ 5.5 was set as the threshold for panel discussion, representing a midpoint between scores of 5 ("somewhat agree") and 6 ("agree"); when some experts assigned lower scores, the resulting lower CI bound tended to fall near 5.5, indicating that the item warranted structured re-evaluation. Items falling between these two thresholds were subject to panel discussion rather than automatic acceptance or rejection.

Results

Literature review and initial identification of candidate indicators

A total of 2,066 articles were reviewed following the systematic literature search (Fig. 1), and 161 studies were selected according to predefined inclusion and exclusion criteria. Preliminary candidate indicators were extracted from the included studies. In total, 59 specific indicators were identified across seven domains: antibiotic use ($n=7$), clinical outcome ($n=8$), microbiological ($n=7$), economic ($n=7$), process-related ($n=8$), department-specific ($n=14$), and other indicators ($n=8$) (Additional file 2).

Refinement of indicators

A total of 17 of 18 (94.4%) panel members responded to the first survey. Based on the free-text responses, panelists proposed revisions, additions, and subdivisions of the preliminary candidate indicators. Subindicators within each major category were further refined and supplemented through this process, and consolidated feedback was shared during an online meeting. As a result of the first-round survey, the initial set of 59 candidate indicators was expanded to 113 indicators, reflecting the addition of newly proposed items and the subdivision of existing indicators into more granular subindicators across all seven domains. All panelists (18/18, 100%) completed the structured questionnaire (second

Delphi round, Additional file 3), and an online expert meeting was convened to review and clarify the survey results, with 13 of the 18 panelists (72.2%) participating. Through these procedures, several indicators from the process-related, department-specific, clinical outcome, microbiological, and economic categories were excluded because of limited applicability, feasibility, interpretability, or measurement challenges, whereas the CDI rate and implementation of therapeutic drug monitoring were retained. In contrast, antibiotic consumption indicators and the ASP core element implementation rate were prioritised for further refinement, and clinician satisfaction was advanced for additional consideration. Consequently, 113 indicators remained, including antibiotic use ($n=23$), clinical outcome ($n=19$), microbiological ($n=9$), economic ($n=11$), process-related ($n=21$), department-specific ($n=19$), and other indicators ($n=11$) (Additional file 4).

Final consensus indicator

The third-round Delphi survey included all 18 panelists. Based on the consensus, a final set of indicators for evaluating ASP effectiveness was established (Table 1). The indicators were organized into four categories: antibiotic use, process-related, department-specific, or other indicators. Consensus was not achieved for the clinical outcome, economic, and microbiological domains.

In the antibiotic use category, six indicators were finalized: (1) antibiotic consumption measured as DOT per 1,000 patient-days; (2) DDD per 1,000 patient-days; (3) reduction rate of broad-spectrum antibiotic use (e.g., carbapenems); (4) rate of intravenous-to-oral antibiotic conversion (e.g., fluoroquinolones); (5) reduction rate of unnecessary combination therapy (e.g., overlapping anaerobic coverage); and (6) change in antibiotic prescription ratio according to the World Health Organization Access, Watch, and Reserve (WHO AWaRe) classification. All indicators were considered suitable for short-term evaluation except for DDD per 1,000 patient-days, which was classified as short- to mid-term. This distinction reflects the inherent methodological complexity of DDD as a metric: because DDD is based on the WHO-defined average adult maintenance dose, institutions with a higher proportion of pediatric or dialysis patients may systematically yield lower DDD values, making short-term cross-institutional comparisons more challenging. In contrast, DOT counts the number of days a patient receives any dose of a given antibiotic regardless of the amount, and is therefore less susceptible to such case-mix confounding, allowing for more immediate interpretation of changes in antibiotic consumption.

In the process-related category, three indicators were selected: (1) appropriate rate of preoperative surgical antibiotic prophylaxis; (2) implementation of therapeutic

Table 1 Outcome indicators related to antimicrobial stewardship programs in Korea: results of the third Delphi round

Indicator	Expected ASP outcome			Feasibility in Korean hospitals			Need for ASP evidence generation			Timeframe	Consensus
	CVR	95% CI	95% CI	CVR	95% CI	95% CI	CVR	95% CI	95% CI		
		Lower	Upper		Lower	Upper		Lower	Upper		
Antibiotic use indicators											
Days of therapy (DOT)/patient days	0.875	6.17	6.83	0.875	6.17	6.83	0.800	6.07	6.93	Short	Accept
Defined daily dose (DDD)/patient days	0.714	6.00	6.78	0.500	5.38	6.50	0.714	5.89	6.78	Short/Mid	Accept
Reduction rate in broad-spectrum antibiotic use (e.g., carbapenem)	0.500	5.43	6.46	0.800	5.82	6.63	0.714	5.57	6.76	Short	Accept
Restricted antibiotic use	0.000	4.57	6.09	0.500	4.88	6.34	0.364	4.85	6.37	Short	Reject
IV-to-oral antibiotic conversion rate (e.g., fluoroquinolones)	0.200	5.31	6.14	−2.500	3.84	5.16	0.875	5.45	6.55	Short	Accept
Inappropriate antibiotic prescription rate (by syndrome or time point)	0.200	5.23	6.22	−16.000	2.63	3.93	0.364	4.52	6.03	Short	Reject
Appropriateness of antibiotic duration (limited to representative syndromes)	0.000	5.12	6.22	−4.000	3.12	4.55	0.200	4.54	6.02	Mid	Reject
Unnecessary combination therapy (e.g., anti-anaerobic duplication)	0.714	5.52	6.59	−0.571	4.39	5.72	0.714	4.95	6.50	Short	Accept
Antibiotic use by WHO AWaRe classification	0.200	4.95	6.05	0.200	4.49	5.96	0.500	4.50	6.05	Short	Accept
Clinical outcome indicators											
Length of stay	−7.000	3.35	4.65	0.364	5.07	6.26	−1.600	4.11	5.33	Long	Reject
ICU length of stay	−16.000	2.74	3.93	0.000	4.87	6.02	−7.000	3.48	4.74	N/A	Reject
ICU admission rate during hospitalization	−16.000	2.70	3.97	0.000	4.77	6.01	−7.000	3.33	4.67	N/A	Reject
C. difficile infection rate†	0.500	5.33	6.34	−0.571	4.41	5.70	0.364	5.32	6.46	Mid	Reject
Infection-related mortality	−16.000	3.59	4.74	−7.000	2.96	4.27	−1.600	3.98	5.36	Short	Reject
Length of stay (limited to specific infection syndromes)	−4.000	4.16	5.06	−0.571	3.74	5.48	−2.500	3.96	5.27	Long	Reject
Adverse drug reaction rate related to antibiotics	−16.000	3.43	4.57	−7.000	3.10	4.56	−1.600	3.68	5.10	N/A	Reject
Surgical site infection rate	−2.500	3.35	4.98	−2.500	3.85	5.04	−1.600	3.23	4.89	N/A	Reject
Acute upper respiratory infection antibiotic prescription rate	−	−	−	0.500	4.96	6.27	−0.250	4.32	5.79	Short	Reject
Economic indicators											
Annual total antibiotic costs (excluding antifungals)	−4.000	3.07	6.26	0.000	4.25	6.98	−1.000	3.47	6.42	Short/Mid	Reject
Antibiotic cost per patient	−2.500	3.59	6.52	−0.250	3.99	6.79	−0.571	3.62	6.49	Mid	Reject
Antibiotic cost per hospital day	−4.000	3.07	6.26	−0.250	3.92	6.75	−0.571	3.50	6.39	Mid	Reject
Cost of operating ASP (limited to personnel expenses)	−7.000	1.71	5.40	0.200	3.57	6.65	−1.600	2.02	5.32	Short	Reject
Cost savings from IV-to-oral conversion	0.000	4.02	6.87	−2.500	2.92	6.08	−1.000	3.78	6.66	Short	Reject
Process indicators											
Appropriate preoperative prophylactic antibiotic use rate	0.800	4.65	7.69	0.364	4.16	7.06	0.714	4.57	7.54	Short	Accept
TDM implementation and dose intervention rate	0.800	4.67	7.66	0.364	3.98	6.91	0.500	4.47	7.30	Short	Accept
Presence of institutional antibiotic use guidelines	0.500	4.53	7.36	0.500	4.53	7.36	0.500	4.32	7.24	Short	Accept

Table 1 (continued)

Indicator	Expected ASP outcome			Feasibility in Korean hospitals			Need for ASP evidence generation			Timeframe	Consensus
	CVR	95% CI Lower	95% CI Upper	CVR	95% CI Lower	95% CI Upper	CVR	95% CI Lower	95% CI Upper		
Intervention rate for prolonged antibiotic prescriptions	0.500	4.46	7.32	−1.000	3.35	6.31	0.200	3.93	6.85	Short	Reject
Intervention rate for combination (duplicate) therapy	0.714	4.33	7.33	−0.571	3.43	6.35	0.364	3.91	6.86	Short	Reject
Concordance of antibiotic use with organism identification	0.800	4.72	7.73	−7.000	2.68	6.21	0.364	4.15	7.08	Short	Reject
<i>Department-specific indicators</i>											
(OR/Surgery) appropriate rate of preoperative surgical antibiotic prophylaxis	0.800	5.82	6.63	0.714	5.47	6.30	0.500	4.85	6.37	Short	Accept
(OR/Surgery) Appropriate timing of postoperative antibiotic discontinuation	0.800	5.71	6.62	0.615	5.10	6.24	0.500	4.69	6.31	Short	Accept
(OR/Surgery) Appropriateness of prophylactic antibiotic selection	0.800	5.82	6.63	0.714	5.57	6.32	0.500	4.74	6.26	Mid	Accept
(ED) Appropriate empirical antibiotic selection for major infections (e.g., UTI, pneumonia)	−0.571	4.54	5.68	−1.000	3.27	4.95	−0.250	4.57	5.76	Mid	Reject
<i>Other indicators</i>											
Number of ASP training sessions conducted	0.200	5.09	6.35	0.800	5.86	6.69	0.200	4.70	6.19	Short	Reject
Use of ASP education indicators [e.g., (training hours × attendance)/staff]	0.200	4.79	6.32	0.714	5.68	6.65	0.200	4.57	6.09	Short	Reject
Healthcare provider satisfaction	−1.600	3.96	5.59	−2.500	3.64	5.02	−1.000	4.47	5.64	Short	Reject
Core elements implementation rate	0.500	5.28	6.61	0.364	4.77	6.23	0.615	5.14	6.53	Short	Accept

Abbreviations: ASP, Antimicrobial Stewardship program; CVR, Content validity ratio; CI, Confidence interval; DOT, Days of therapy; DDD, Defined daily dose; TDM, Therapeutic drug monitoring; ICU, Intensive care unit; ED, Emergency department; OR, Operating room; UTI, Urinary tract infection; WHO, World Health Organization; AWaRe, Access, watch, and reserve.

Timeframe: Short = 3–4 years; Mid = 5–9 years; Long = ≥10 years; N/A = Not applicable.

† *C. difficile* infection rate: Expected ASP Outcome CVR met the threshold (0.500); however, Feasibility CVR did not meet the acceptance criterion (CVR = −0.571). The Reject judgment was retained.

‡ (ED) Appropriate empirical antibiotic selection: Not adopted primarily due to concerns regarding operational feasibility (systematic tracking not yet widely available in Korean hospitals), case heterogeneity, and reliance on microbiological confirmation data unavailable at the time of ED prescription.

drug monitoring and dose adjustment recommendations; and (3) establishment of institutional antibiotic use guidelines. All were classified as short-term evaluation measures.

In the department-specific category, three indicators—primarily related to surgical and perioperative care—were finalized: (1) appropriate rate of preoperative surgical antibiotic prophylaxis; (2) appropriateness of antibiotic discontinuation timing after surgery; and (3) appropriateness of antibiotic selection for surgical prophylaxis. The first two were classified as short-term evaluation indicators, whereas the latter was considered mid-term.

Finally, in the other indicators category, the implementation rate of core ASP elements—defined as a standardized set of structural and organizational components essential for establishing effective ASPs, including

leadership commitment, accountability, drug expertise, tracking, reporting, and education [7]—was identified as a comprehensive structural indicator reflecting overall ASP performance and was considered appropriate for short-term assessment.

Discussion

This study represents the first national consensus in Korea on ASP performance indicators developed through a rigorous Delphi process in the context of a newly launched government-led reimbursement pilot. Through a systematic evidence review and structured expert consensus, we identified seven major domains and initially defined 59 indicators encompassing both short-term measures (antibiotic consumption and adherence to restricted antibiotic use) and medium-to long-term

measures (resistance trends and clinical outcomes). Through the subsequent Delphi process, this list was iteratively refined to a final set of 13 core indicators comprising six antibiotic use metrics, three process-related indicators, three indicators targeting high-risk clinical settings, and one additional cross-cutting measure. Rather than replicating earlier single-center or drug-class-focused analyses, this framework was designed to enable nationwide benchmarking at a critical juncture when ASPs are formally integrated into national health policy.

In Korea, antibiotic consumption remains among the highest compared with other high-income countries, and national strategies have increasingly prioritised ASPs [10]. Although many hospitals have independently implemented ASPs, these efforts have largely remained confined to individual institutions. Previous single-center studies have demonstrated reductions in antibiotic use and resistance through infectious disease specialist-led or stepwise ASP models, confirming the feasibility of institutional ASP implementation [12–14]. However, such experiences cannot substitute for a harmonized, evidence-based national framework. Therefore, the introduction of a reimbursement-based ASP pilot program in 2024 provided a timely opportunity to develop standardized indicators that could ensure consistent evaluation and support sustainable policy expansion.

Notably, the indicators that achieved consensus were overwhelmingly structural and process-oriented, including antibiotic consumption metrics (DOT, DDD, and broad-spectrum reduction), intravenous-to-oral conversion, appropriateness of surgical prophylaxis, and implementation of core ASP elements. These domains are broadly consistent with established ASP indicator frameworks from other countries, which similarly prioritise antibiotic use monitoring, prescribing processes, and stewardship infrastructure [18, 19]. However, unlike some international frameworks, no clinical outcome, economic, or microbiological indicators met the predefined validity thresholds in our study. This distribution reflects a pattern similar to international experiences in which process measures consistently dominate ASP evaluation frameworks. For instance, O’Riordan et al. reported that 75% of 229 indicators across 16 studies were process-related measures, whereas only 1% were outcome indicators [18]. Similarly, the ASP indicator set established by Centers for Disease Control and Prevention is composed almost entirely of structure and process-related measures, with no routine requirements for clinical outcomes [19]. These differences likely reflect practical and contextual challenges in standardising indicators that are highly dependent on institutional and healthcare system characteristics. Clinical and microbiological outcome measurement is often hampered by

variability in diagnostic criteria, laboratory capacity, surveillance practices, reporting standards, and patient case mix [20]. Consistent with these challenges, our Delphi panel found it difficult to reach consensus on CDI incidence as a clinical outcome because of the heterogeneity in diagnostic methods and surveillance practices across hospitals in Korea. Panel discussions revealed that many hospitals remain in the early stages of ASP implementation, and the wide variability in baseline resistance, patient populations, and stewardship interventions makes the derivation of standardized outcome indicators currently impractical. As ASP infrastructure and data systems advance, selected outcome measures may be incorporated, particularly in settings with standardized diagnostic practices and surveillance systems. However, as demonstrated in international frameworks, outcome indicators remain difficult to operationalize and are often included only conditionally or on a limited basis [21]. Thus, any future expansion of outcome metrics should be pursued cautiously and only when data comparability and methodological robustness are ensured.

From a policy perspective, the indicators identified through the Delphi process—primarily antibiotic-use metrics, process-related indicators such as intravenous-to-oral conversion and surgical prophylaxis appropriateness, and structural elements aligned with core ASP activities—were deliberately selected for feasibility within existing hospital information systems, enabling near-term implementation. At the national level, these metrics can facilitate benchmarking, feedback mechanisms, and performance-based reimbursement [19, 21, 22]. Importantly, health authorities can use the indicator set to systematically identify areas requiring improvement, inform priority setting, and support evidence-based policy formulation, as ASP implementation expands nationwide. Although the framework was developed specifically for the Korean healthcare system, some of the principles identified through this consensus process may be informative for other settings facing similar implementation and resource challenges. Furthermore, defining a concise core indicator set for smaller communities or long-term care hospitals is essential to ensure equitable participation in future national programs. Successful implementation will require further operational refinement of several indicators that have reached consensus but lacked standardized definitions [21, 22]. In particular, metrics such as broad-spectrum antibiotic reduction and intravenous-to-oral conversion, require clearer specification of target drug classes and eligibility criteria to ensure consistent measurement across hospitals. Early assessments may focus on overall changes in antibiotic use; however, long-term approaches should stratify indicators by disease category, clinical department, or patient risk group to enhance interpretability and actionability.

This study had two main limitations. First, although our consensus was grounded in systematic evidence and expert opinion, empirical validation of indicator reliability and responsiveness remains necessary [9, 18]. Second, although we conducted a structured literature review to identify candidate indicators, some relevant indicators may not have been captured because of limitations in search strategy, database coverage, or reporting in the published literature. In addition, the final indicator set may have been influenced by the composition of the expert panel and the operational context of Korean hospitals. The limited inclusion of clinical outcome indicators likely reflects both the early stage of ASP implementation in Korea and the methodological challenges of standardizing outcome attribution across heterogeneous settings. Future research should incorporate these dimensions to assess the long-term impact of ASP performance and establish the association between process measures and patient outcomes.

Conclusions

This study provides Korea's first national consensus on ASP performance metrics at a critical policy juncture. As the first nationally scaled effort to standardize ASP evaluation, the proposed framework addresses the urgent needs identified by the 2024 government-led reimbursement pilot program. Although tailored to the Korean healthcare system, these indicators may serve as valuable references for comparable healthcare environments with limited resources. Adoption of these standardized metrics can enhance ASP accountability, guide equitable reimbursement, support the maturation of institutional ASPs, and foster international collaboration in evaluating ASP effectiveness.

Abbreviations

ASP	Antimicrobial stewardship program
AWaRe	Access, watch, and reserve (World Health Organization classification)
CDI	Clostridioides difficile infection
CVR	Content validity ratio
DDD	Defined daily dose
DOT	Days of therapy
QI	Quality indicator

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13756-026-01778-9>.

Supplementary Material 1.

Acknowledgements

The authors gratefully acknowledge Hyunok Yun for her technical support in systematic literature extraction as well as all individuals who contributed their expertise to the Delphi survey.

Author contributions

Conceptualization: Kim B and Kim HB. Data curation: Park SY, Lim SM, Lee MJ and Lee R. Formal analysis: Park SY, Kim J, Yoon C and Park JY. Investigation: All authors. Methodology: Lim SM, Lee R, Park SY, and Kim B. Project administration: Kim B. Resources: All authors. Software: Park SY, Kim YC and Lim SM. Supervision: Moon SM, Kim HB, Park SY, Kim YC and Kim B. Validation: Lim SM, Lee R and Moon SM; Visualization: Lee R, Park SY and Kim B; Writing—original draft: Lim SM, Lee R, Kim B and Park SY; Writing—review and editing: All authors. All authors read and approved the final manuscript.

Funding

This work was supported by the Research Program funded by the Korea Disease Control and Prevention Agency (Grant No: 2024-10-012). The funder had no role in the study design, data collection, analysis, manuscript preparation, or decision to publish.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Institutional Review Board (IRB) of Hanyang University Seoul Hospital (IRB number: 2025-02-015). The requirement for written informed consent from the participants was waived.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

- ¹Department of Pediatrics, Hanyang University College of Medicine, Seoul, Republic of Korea
- ²Division of Infectious Diseases, Department of Internal Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
- ³Division of Infectious Diseases, Department of Internal Medicine, Hanyang University Seoul Hospital, Seoul, Republic of Korea
- ⁴Department of Internal Medicine, Hanyang University College of Medicine, Seoul, Republic of Korea
- ⁵Department of Pediatrics, Severance Children's Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea
- ⁶Department of Internal Medicine, Inje University Sanggye-Paik Hospital, Inje University College of Medicine, Seoul, Republic of Korea
- ⁷Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Republic of Korea
- ⁸Department of Internal Medicine, Yonsei Severance Hospital, Yonsei University College of Medicine, Yonsei, Republic of Korea

Received: 22 January 2026 / Accepted: 5 June 2026

Published online: 15 June 2026

References

1. O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. Review on Antimicrobial Resistance; 2016.
2. World Health Organization. Global action plan on antimicrobial resistance. Geneva: World Health Organization; 2015.
3. Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, et al. Implementing an antibiotic stewardship program: guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis. 2016;62(10):e51–77. <https://doi.org/10.1093/cid/ciw118>.
4. Schuts EC, Hulscher MEJL, Mouton JW, Verduin CM, Stuart JWTC, Overdiek HWPM, et al. Current evidence on hospital antimicrobial stewardship objectives: a systematic review and meta-analysis. Lancet Infect Dis. 2016;16(7):847–56. [https://doi.org/10.1016/S1473-3099\(16\)00065-7](https://doi.org/10.1016/S1473-3099(16)00065-7).

5. Dalton BR, Sabuda DM, Bresee LC, Conly JM. Assessment of antimicrobial utilization metrics: days of therapy versus defined daily doses and pharmacy dispensing records versus nursing administration data. *Infect Control Hosp Epidemiol*. 2015;36(6):688–94. <https://doi.org/10.1017/ice.2015.46>.
6. van den Bosch CMA, Hulscher MEJL, Natsch S, Gyssens IC, Prins JM, Kullberg BJ, et al. Development of quality indicators for appropriate antibiotic use in adults in Dutch hospitals. *J Antimicrob Chemother*. 2018;73(6):1611–9. <https://doi.org/10.1093/jac/dky055>.
7. Cheong HS, Park K-H, Kim HB, Kim S-W, Kim B, Moon C, et al. Core elements for implementing antimicrobial stewardship programs in Korean general hospitals. *Infect Chemother*. 2022;54(4):637–73. <https://doi.org/10.3947/ic.2022.0171>.
8. Moehring RW, Anderson DJ, Cochran RL, Hicks LA, Srinivasan A, Dodds Ashley ES, et al. Expert consensus on metrics to assess the impact of patient-level antimicrobial stewardship interventions in acute-care settings. *Clin Infect Dis*. 2017;64(4):377–83. <https://doi.org/10.1093/cid/ciw787>.
9. Schweitzer VA, van Heijl I, van Werkhoven CH, Islam J, Hendriks-Spoor KD, Bielicki J, et al. The quality of studies evaluating antimicrobial stewardship interventions: a systematic review. *Clin Microbiol Infect*. 2019;25(5):555–61. <https://doi.org/10.1016/j.cmi.2018.11.002>.
10. Yun IJ, Park HJ, Chae J, Heo SJ, Kim YC, Kim B, et al. Nationwide analysis of antimicrobial prescription in Korean hospitals between 2018 and 2021: the 2023 KONAS report. *Infect Chemother*. 2024;56(2):256–65. <https://doi.org/10.3947/ic.2024.0013>.
11. Jeong YI, Lee HY, Lee S, Jeong GY, Kim SH, Kim S, et al. Korea's national action plan on antimicrobial resistance: focusing on the appropriate use of antibiotics. *Infect Chemother*. 2025;57(1):203–14. <https://doi.org/10.3947/ic.2025.0028>.
12. Hwang H, Kim B. Impact of an infectious diseases specialist-led antimicrobial stewardship programme on antibiotic use and antimicrobial resistance in a large Korean hospital. *Sci Rep*. 2018;8(1):14757. <https://doi.org/10.1038/s41598-018-33201-8>.
13. Lee SJ, Lee R, Cho SY, Nho D, Ahn HL, Lee DG. Implementing antimicrobial stewardship: lessons and perspectives from a university-affiliated tertiary hospital in Korea. *Korean J Intern Med*. 2024;39(3):399–412. <https://doi.org/10.3904/kjim.2023.504>.
14. Shin DH, Kim HS, Heo E, Shin MJ, Kim NH, Lee H, et al. Stepwise expansion of antimicrobial stewardship programs and its impact on antibiotic use and resistance rates at a tertiary care hospital in Korea. *Microbiol Spectr*. 2022;10(1):e0033522. <https://doi.org/10.1128/spectrum.00335-22>.
15. Lee H, Moon SM, Song K, et al. National pilot program for antimicrobial stewardship in the Republic of Korea. *JAMA Netw Open*. 2026;9(1):e2555843. <https://doi.org/10.1001/jamanetworkopen.2025.55843>.
16. Lawshe CH. A quantitative approach to content validity. *Pers Psychol*. 1975;28(4):563–75. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>.
17. Ayre C, Scally AJ. Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Meas Eval Couns Dev*. 2014;47(1):79–86. <https://doi.org/10.1177/0748175613513808>.
18. O'Riordan F, Shiely F, Byrne S, Fleming A. Quality indicators for hospital antimicrobial stewardship programmes: a systematic review. *J Antimicrob Chemother*. 2021;76(6):1406–19. <https://doi.org/10.1093/jac/dkab034>.
19. Pollack LA, Plachouras D, Sinkowitz-Cochran R, Gruhler H, Monnet DL, Weber JT, et al. A concise set of structure and process indicators to assess and compare antimicrobial stewardship programs among EU and US hospitals: results from a multinational expert panel. *Infect Control Hosp Epidemiol*. 2016;37(10):1201–11. <https://doi.org/10.1017/ice.2016.115>.
20. Stemkens R, Schouten JA, van Kessel SAM, Akkermans RP, Telgt DSC, Fleuren HWWA, et al. How to use quality indicators for antimicrobial stewardship in your hospital: a practical example on outpatient parenteral antimicrobial therapy. *Clin Microbiol Infect*. 2023;29(2):182–7. <https://doi.org/10.1016/j.cmi.2022.07.007>.
21. Abdel Hadi H, Eltayeb F, Al Balushi S, Daghfal J, Ahmed F, Mateus C. Evaluation of hospital antimicrobial stewardship programs: implementation, process, impact, and outcomes, review of systematic reviews. *Antibiot (Basel)*. 2024;13(3):253. <https://doi.org/10.3390/antibiotics13030253>.
22. Kim T, Moehring RW, Ashley ED, Kim HS, Anderson DJ. Strengthening antimicrobial stewardship in Korea: strategies to address workforce gaps and expand ASP nationwide. *Infect Chemother*. 2025;57(1):185–93. <https://doi.org/10.3947/ic.2024.0135>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.