**Supplemental Table 1**

**Glossary**

**Clinical staff** – the test utilization management practices assessed in this review were generally directed to clinical staff authorized to order laboratory testing, as depicted in Supplemental Tables 4 and 5. Because included studies occurred both within the US and outside of the US, terminology used varies, including terminology that may be unfamiliar to some readers.

Attending physicians – physician staff who have completed residency and practice medicine. They are sometimes referred to as faculty, or consultants.

Clinician – general term including physicians and non-physicians (e.g. physician assistants, nurse practitioners, and others if mentioned).

Emergency physician – physician who specializes in emergency medicine.

Hospitalist – refers to hospital physicians.

House staff/ officers – physician staff who are interns, residents, fellows, and medical registrars. They are sometimes referred to as physicians in training.

Medical students – students who are working towards a medical degree.

Ordering clinician/provider – includes all groups who are authorized to order lab tests for patients, such as attending’s, house staff, physician assistants, nurse practitioners, and others.

Primary care physicians – refers to general practitioners (GPs), family practitioners, internists.

**Cohen’s *d*** – an effect measure that can be derived when groups are compared on continuous outcomes, and represents a standardized difference between the means of two groups relative to the pooled standard deviation. Statistical techniques are available to convert between Cohen’s *d* and OR effect measures.

**Combined Practices** – various combinations of practices may be implemented in support of appropriate test utilization. The range of possible combinations is extensive, as is the range of timing, frequency, or intensity of combined components.

**Computerized decision support system/tools (CDSS/CDST)** – process for enhancing health-related decisions and actions through integration of pertinent clinical knowledge and patient information, as may support appropriate test utilization by offering clinician’s patient-specific diagnostic testing or treatment advice or alerts. Within this systematic review the following specific interventions were encountered, grouped as CDSS/CDST: systems that propose guideline-based investigations for indications/suspected conditions inputted by the clinician, in relation to patient-specific indications inputted by the clinician; and systems that provide testing guideline alerts in relation to patient-specific data (demographics and/or other available test results).

**Computerized provider order entry (CPOE)** – computerized provider order entry systems automate laboratory ordering processes and ensure standardized, legible, and complete orders. These systems have the potential to automate the clinical test ordering process and to improve the quality and safety of patient care, with practice interventions representing new CPOE systems replacing written test orders, or modifications to existing CPOE systems. In the latter case, this systematic review encountered the following specific interventions: redundant test alerts in relation to appropriate test-specific intervals for testing; limits on advance ordering of testing; display of test charges; limiting test availability with the CPOE interface; and structuring CPOE according to universal ordering routines.

**Consistency** – when population parameter effects are the same across the included studies.

**Diagnostic Management Teams** – provides expert, multidisciplinary consult services and support to providers in the diagnostic pathway, guiding individual diagnostic evaluations and interpretations in real-time, through aggregation of clinical information and test information and generation of patient-specific narratives. They provide on-demand advice for ordering appropriate tests for individual patients (or related groups of patient cases) and interpreting and acting on test information, in order to reduce diagnostic errors and improve patient outcomes while supporting healthcare savings. DMTs, therefore, have the potential to impact inappropriate test utilization (over- and/or under-utilization of necessary testing) within a healthcare organization.

**Education** – education and continuing medical education consists of activities explicitly serving to maintain, develop, or increase the knowledge, skills and professional performance in relation to appropriate test utilization.

**Effect size** – the association between two or more study’s outcome measures for the group in which the intervention/practice was evaluated and those of its control or comparison group. Effect size ratings can be numeric or reflect the magnitude of effect in qualitative terms. Numeric representation can be converted to qualitative values through expert consensus on cutoffs representing a “substantial”, “moderate”, and “minimal” effect; however, setting cutoffs for qualitative ratings invariably has an element of subjective judgment.

**External validity, generalizability, applicability** – extent to which the effects observed in the study are applicable to populations other than the study’s population.

**Facility setting** – an aspect of applicability or generalizability, facility setting is categorized and defined in this systematic review as follows, with more than one category often applying (e.g. a hospital as both academic and children’s), and with placement of individual studies in a category generally based on how authors characterized their facility. The following web resources informed definitions: American Hospital Association ([www.aha.org](http://www.aha.org)), Becker’s Hospital Review ([www.beckershospitalreview.com](http://www.beckershospitalreview.com)), Lab Tests Online (<https://labtestsonline.org>), and the Association of American Medical Colleges (www.aamc.org).

Children’s Hospital – a hospital whose patient population is exclusively pediatric, frequently from birth to < 18 years of age.

Community Hospital – the American Hospital Association (AHA) broadly characterizes community hospitals as a non-federal, short-term general hospitals and can include specialty hospitals (women’s, eye, orthopedic, rehabilitation, heart, etc.).

International – additional category designation for studies occurring outside of the United States. Healthcare policies, structures, processes, and payment models frequently vary in other countries, forming the context in which the delivery of care occurs, and potentially impacting generalizability of findings.

Not otherwise specified – category applies when facility is a medical facility serving patients for whom laboratory testing is ordered, but is otherwise not characterized by study authors.

Primary Care Offices/Outpatient Clinics – outpatient services involving basic or general health care traditionally provided by doctors trained in family practice, pediatrics, internal medicine, and occasionally gynecology. This category also includes urgent care centers.

Public Hospital – a hospital having government ownership, federal, state, or local (county/city). Public hospitals are generally open to the general public, with ability to care for patients unable to pay. When including VA hospitals in this group, public hospitals comprise approximately 25% of hospitals nationwide.

Reference Laboratory – private, commercial clinical laboratories independent of a specific healthcare organization, and which perform testing referred by a variety of healthcare settings (e.g. hospitals, clinics, etc.).

University/Academic/Teaching Hospital – a hospital having association with at minimum with a medical school and residency training program, and may in addition include association with other health professions schools (e.g. clinical laboratory science programs, nursing programs, etc.).

Veterans Affairs (VA) Hospital – a federal government hospital serving military veterans. VA hospitals represent the largest federal hospital system, with approximately 150 hospitals nationwide.

**Feedback** – feedback presents information on test ordering patterns back to the users, individually or in group settings, in relation to appropriate utilization criteria. The effects of formal assessment and feedback on physician performance may be influenced by the source and duration of feedback, and whether comparative performance is visible to peers.

**Forest plot** – a graphic representation of effect sizes for two or more studies included in a meta-analysis

**Heterogeneity** – the state in which there is between-study variation in effect sizes in the population

***I*2** – a standardized statistic which quantifies the percent of total variability due to heterogeneity, and is expressed as a percentage between 0% (no variability due to heterogeneity) and 100% (all variability due to heterogeneity).

**Internal validity** – extent to which the design and conduct of the study are likely to prevent systematic error. Internal validity is a prerequisite for external validity.

**Laboratory Test Utilization Team** – a multidisciplinary task force providing recommendations on systematic efforts within a healthcare organization, and guidance on setting standards and expectations for ongoing utilization management, introduction of testing, and elimination of testing. It may recommend test utilization protocols within the organization, it may establish criteria for defining occurrences of “inappropriate” testing, or it may promote/monitor uptake of systematic mechanisms to support appropriate test utilization (e.g., CDSS, education, feedback, reflex testing, etc.). Additionally they may oversee efforts to validate test utilization criteria in the local healthcare settings, as well as efforts to systematically audit test utilization and gather/analyze utilization data. LTUs are also called clinical laboratory utilization committees, clinical laboratory advisory committees, and laboratory diagnostic committees.

**Meta-analysis** – the process of using statistical methods to standardize and quantitatively combine the results of similar studies in an attempt to allow inferences to be made from a collection of studies. It allows for estimates of effects across studies. **Odds ratio** – an effect measure that can be derived when groups are compared on dichotomous outcomes, and represents the odds of an event/outcome occurring in one group in relation to the odds of it occurring in another group. Statistical techniques are available to convert between Cohen’s *d* and *OR* effect measures, to facilitate more robust comparison among studies. In comparison of numbers of tests, all results were expressed as *OR*s.

**Reflex Testing** – a reflexed test is any test that automatically results in the order of one or more secondary tests based on preset criteria applied to the initial test, in support of appropriate test utilization.

**Systematic review** – a scientific investigation that focuses on a specific question and uses explicit, planned scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may or may not include a quantitative synthesis of the results from separate studies (meta-analysis).

**Test Review** – test review involves consultation with ordering physicians on specific laboratory testing at the time it is ordered, for determination whether it should be performed. Test review can be done by laboratory experts, diagnostic management teams, laboratory consultation teams, or test gatekeepers.