REVIEW ARTICLE (STRUCTURED REVIEW WITH QUANTITATIVE EVIDENCE SYNTHESIS)

Quality of medical examination of conscription-age men: a structured review with quantitative evidence synthesis

Shoira Q. Abdusattarova[1*], Saidakhon Sh. Mirrakhimova[1]

[1]*Tashkent, Republic of Uzbekistan*

*Corresponding author: Shoira Q. Abdusattarova; e-mail: nagaeva.gulnora@mail.ru

Abstract

Background. Medical examination of conscription-age men is performed largely by civilian primary care, yet its quality is judged almost exclusively by downstream military outcomes.

Aim. To quantify the evidence on disease burden in this population, on the diagnostic performance of the components of the examination, and on interventions that change professional performance, and to derive a testable indicator framework for quality management.

Methods. Structured review of PubMed/MEDLINE and Cochrane sources to October 2026 (meta-analyses, randomised controlled trials [RCTs], national cohorts and regulatory documents of the Republic of Uzbekistan). Estimates were extracted as reported; where counts or confidence intervals permitted, proportions with Wilson 95% CI, odds/risk ratios with Woolf 95% CI and exact two-sided p values were computed (χ^2 or Fisher's exact test; Altman–Bland method).

Results. Prevalence in conscription-age men spanned more than two orders of magnitude, from 23.5% for myopia to 0.12% for HIV, with potentially lethal cardiac disorders in 0.34% (95% CI 0.29–0.40)†. Obesity among 1.5 million Austrian conscripts rose from 1.6% to 8.2% between 1983 and 2017 (+6.6 percentage points; $p < 0.001$†). Within the examination, diagnostic performance differed sharply: electrocardiography detected 94% of lethal cardiac disorders against 20% for history and 9% for physical examination. Concordance between records and self-report was no better than moderate (κ 0.23–0.51), so measured prevalence depended on the method of ascertainment. Audit and feedback, the only quality intervention with large randomised evidence (292 RCTs), produced a mean absolute improvement in desired practice of 6.2% (95% CI 4.1–8.2; OR 1.47, 1.31–1.64), rising when baseline performance was low. Regional fitness for service rose by 14.4 percentage points over 12 years (OR 2.02; 1.94–2.11; $p < 0.001$†) where health-improvement work was organised systematically.

Conclusion. The examination is currently managed as a legal procedure and not as a measurable clinical process. A 15-indicator structure–process–outcome framework, with instrumentally verified rather than self-reported inputs and audit and feedback as the delivery mechanism, is proposed for prospective testing. † Calculated by the authors.

Keywords: *conscript; medical examination; military medical expertise; primary health care; quality indicator; quality of health care.*

MAIN TEXT

Introduction

Compulsory military service brings an entire male birth cohort into contact with the health system at a single, legally fixed moment. In the Republic of Uzbekistan this contact is governed by the Regulation on medical examination in the Armed Forces, approved by Presidential Decree PD-4076 of 25 December 2018, and by Decree PD-4630 of 6 March 2020 on the conscription procedure; the underlying duty is established by the Law on universal military duty and military service. The stated purpose of the Regulation is the improvement of the quality and effectiveness of the examination and the staffing of the Armed Forces with healthy and physically developed servicemen.

The decisive stages of that examination are performed by civilian primary care. Territorial health bodies and outpatient facilities carry out annual examinations of males aged 15–16 years, mandatory diagnostic tests and oral sanitation before the commission, additional outpatient or inpatient examination, treatment before conscription, and the transfer of medical records within ten days; the commissions themselves are staffed by specialists of the state health system. The examination is therefore a clinical process delivered by primary care under a legal deadline, and not a military procedure.

Two bodies of evidence bear on whether that process performs well. The first concerns what there is to find. Conscription-age men are not a healthy reference population: myopia affects 23.5% of Finnish conscripts [1], one in five Finnish conscripts experiences back pain during service [2], obesity among Austrian conscripts rose fivefold in 35 years [3], adult attention-deficit/hyperactivity disorder affects 2.8% of Korean servicemen with extreme psychiatric comorbidity [4], and roughly one in 294 young persons screened carries a potentially lethal cardiac disorder [5]. Conversely, the same examination has served as a national surveillance platform for infection [6]. The second body concerns what the examination can detect. Within a single screening protocol, the diagnostic performance of the components differs by an order of magnitude: electrocardiography detects 94% of lethal cardiac disorders, a structured history 20% and physical examination 9% [5]. What is counted as prevalent therefore depends on which instrument is used, and self-reported and recorded information agree only fairly to moderately (κ 0.23–0.51) [7–9].

Quality of care has been assessed through the triad of structure, process and outcome since Donabedian [10]. Applied to conscription, the Uzbek regulation measures outcome only, and only after the civilian stages are complete: returns from the assembly point and the discharge of wrongly

conscripted servicemen, each triggering a disciplinary inquiry. No indicator describes coverage, completeness or timeliness of the work that produced those outcomes. Whether such measurement would change anything is itself an empirical question with a large randomised literature: audit and feedback has been tested in 292 trials [11]. Studies of conscript health in the post-Soviet space document the burden and the trend in fitness [12–14] but do not treat the examination as a measurable process.

The aim of this review was therefore to quantify (i) the burden of disease in conscription-age men and the consequences of imprecise selection; (ii) the diagnostic performance of the components of the examination and the agreement between its sources of information; (iii) the effect of interventions that change professional performance; and (iv) what the current regulatory pathway actually measures, and to derive from these a testable indicator framework for the management of examination quality.

Material and Methods

Design and information sources

A structured review with quantitative evidence synthesis was performed; reporting followed the applicable items of PRISMA 2020 [15]. PubMed/MEDLINE was searched without language restriction from inception to October 2026, supplemented by the Cochrane Database of Systematic Reviews, by the reference lists of the retrieved reviews, and by the National Database of Legislation of the Republic of Uzbekistan (lex.uz) for the regulatory corpus. Search strings combined population terms (“conscripts”, “military recruits”, “conscription-age”, “adolescent males”) with domain terms (“prevalence”, “medical examination”, “fitness for military service”, “screening”, “sensitivity and specificity”, “attrition”, “audit and feedback”, “quality indicator”, “quality of health care”).

Eligibility

Eligible were systematic reviews and meta-analyses, randomised controlled trials, and national or prospective cohort and case–control studies reporting a quantitative estimate within one of four domains: prevalence or incidence of disease in conscription-age men; diagnostic performance or agreement of the components of a health examination; effect of an organisational intervention on professional performance or on a service outcome; and the association of examination findings with later service outcomes. Regulatory acts in force in the Republic of Uzbekistan were analysed as a separate corpus. Case reports, narrative reviews without data, and studies confined to therapeutic interventions were excluded.

Data extraction and derived statistics

For each estimate, design, number of studies (*k*) and participants (*n*), effect measure, 95% confidence interval (CI) and exact *p* value were extracted as reported, and all numbers were re-checked against the source. Derived statistics are flagged with a dagger (†). Proportions are given with Wilson 95% CI. When 2×2 counts were available or could be reconstructed from reported percentages and denominators, odds ratios with Woolf (logit) 95% CI and two-sided *p* values (Pearson χ^2 for $n>1{,}000$; Fisher's exact test otherwise) were computed, together with risk differences and their normal-approximation CI. When only an estimate and its CI were reported, *p* was obtained by the Altman–Bland method on the log scale for ratio measures [16]. No new pooling across sources was undertaken because of clinical and methodological heterogeneity. Calculations were performed in Python 3 (SciPy 1.17); $\alpha=0.05$ (two-sided). Certainty is reported as rated by the source (GRADE or Cochrane risk of bias). The review used published aggregate data and public regulatory documents only and did not require ethical approval.

Results

Burden of disease and the cost of imprecise selection

Prevalence estimates in conscription-age men spanned more than two orders of magnitude (Figure 1, Table 1). Myopia under cycloplegia affected 23.5% (95% CI 21.3–25.9)† of 1,286 Finnish conscripts, with near-work time (adjusted OR 1.18 per unit; $p<0.001$†) and outdoor time (0.82; $p<0.001$†) as the dominant modifiable correlates [1]. Back pain occurred in 20.0% (18.8–21.3)† of 4,029 Finnish conscripts and was unchanged across 1987–2005; pre-service back problems tripled the risk (RR 2.03; $p=0.002$) and half of all episodes arose in the first months of service [2]. Among 1.5 million Austrian conscripts, obesity rose from 1.6% (1.6–1.7) in 1983 to 8.2% (8.1–8.3) in 2017, an absolute increase of 6.6 percentage points (6.5–6.7; $p<0.001$)† [3].

Table 1. **Burden of disease in conscription-age men and the measurable consequences of imprecise selection**

Domain and condition	Source (design, country)	Evidence base		Outcome	Estimate (95% CI)	*p*	Certainty / remarks
		Studies	Participants				
A. Burden of disease in conscription-age men							
Refractive error	Pärssinen et al., 2025 (cross-sectional, Finland) [1]	1	1,286	Myopia under cycloplegia	23.5% (21.3–25.9)†	—	High myopia 0.4%
				Near-work time, per hour	aOR 1.18 (1.11–1.26)	<0.001†	Outdoor time protective: aOR 0.82 (0.75–0.90)
Musculoskeletal	Suikkanen et al., 2023 (register cohort, Finland) [2]	1	4,029	Back pain during service	20.0% (18.8–21.3)†	—	Stable across 1987–2005; half of episodes in first months
				Back problems reported before service	RR 2.03	0.002	Smoking RR 1.35; basic education only RR 1.55 (both p<0.001)
Metabolic	Yang et al., 2021 (national trend, Austria) [3]	1	1.5 million	Obesity (BMI ≥30), 1983 vs 2017	1.6% (1.6–1.7) → 8.2% (8.1–8.3)	<0.001†	Absolute rise 6.6 pp (6.5–6.7)†; complete conscript cohorts
Mental health	Yeom et al., 2020 (survey, Republic of Korea) [4]	1	3,441	Adult ADHD (A-ADHD scale)	2.8% (2.3–3.4)†	—	Self-report instrument, not clinical interview
				Comorbid social anxiety; suicidality	OR 40.5 (25.1–65.7); 11.0 (6.7–17.9)	<0.001†	Firth logistic regression

Domain and condition	Source (design, country)	Evidence base		Outcome	Estimate (95% CI)	*p*	Certainty / remarks
		Studies	Participants				
Cardiac	Harmon et al., 2015 (SR/MA of screening programmes) [5]	15	47,137	Potentially lethal cardiac disorder	0.34% (0.29–0.40)†	—	1 in 294; WPW 42%, long QT 11%, HCM 11%
Infection	Sperhacke et al., 2018 (national survey, Brazil) [6]	1	37,282	HIV seroprevalence	0.12% (0.09–0.16)†	—	Conscription used as a national surveillance platform
Nutrition	Carrasco et al., 2025 (SR/MA) [17]	22 (6 pooled)	124,414	Anaemia and iron deficiency over training	No significant change	NS	Fair-to-good cross-sectional, poor longitudinal studies
B. Consequences of imprecise selection							
Attrition	Gendler et al., 2023 (cross-sectional, Israel) [18]	1	46,472	Attrition from basic combat training	10.0% (9.7–10.3)†	—	>11 sick-leave days: 59% vs 19% attrition
Long-term health	Crump et al., 2016 (national cohort, Sweden) [19]	1	1,534,425	Low fitness at 18 y → type 2 diabetes	aHR 3.07 (2.88–3.27)	<0.001	39.4 million person-years; holds at normal BMI

† Calculated by the present authors (Wilson 95% CI for proportions; Woolf CI and χ^2 for ratios; Altman–Bland method for p from reported CI). aHR, adjusted hazard ratio; aOR, adjusted odds ratio; BMI, body mass index; NS, not significant; pp, percentage points; RR, risk ratio; SR/MA, systematic review/meta-analysis.

Conditions that are invisible to inspection were equally frequent. Adult ADHD affected 2.8% (2.3–3.4)† of 3,441 Korean servicemen and carried extreme comorbidity (social anxiety OR 40.5, 95% CI 25.1–65.7; suicidality 11.0, 6.7–17.9; both $p<0.001$†) [4]. Potentially lethal cardiac disorders were found in 0.34% (0.29–0.40)† of 47,137 young persons screened, that is one in 294 [5], and HIV in 0.12% (0.09–0.16)† of 37,282 Brazilian conscripts [6]. By contrast, basic training itself did not change the prevalence of anaemia or iron deficiency in a meta-analysis of 22 studies [17].

The consequences of imprecise selection are measurable at both ends. Attrition from basic combat training reached 10.0% (9.7–10.3)† of 46,472 Israeli soldiers and was concentrated in identifiable subgroups: those with more than eleven sick-leave days attrited at 59% against 19% [18]. In the opposite direction, fitness measured at the examination predicted disease decades later: among 1,534,425 Swedish conscripts followed for 39.4 million person-years, the combination of low aerobic capacity and low muscle strength at age 18 carried a threefold risk of type 2 diabetes (adjusted HR 3.07; 2.88–3.27; $p<0.001$), even at normal body mass index [19, 20].

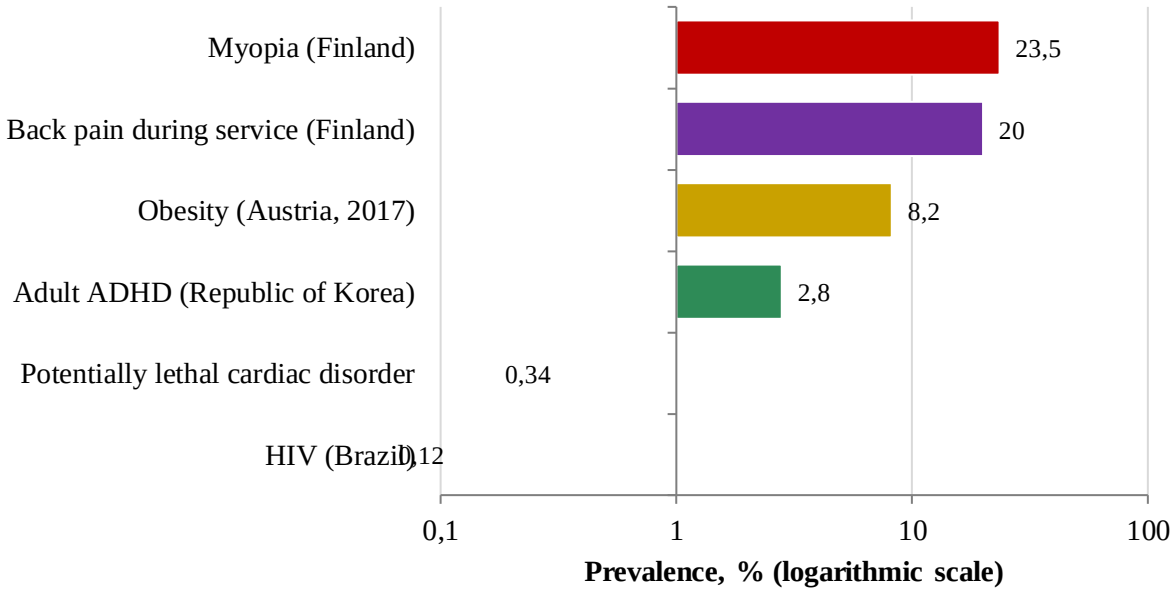


Figure 1. Prevalence of selected conditions in conscription-age men (logarithmic scale). Derived from population and screening studies [1–6]; the estimate for lethal cardiac disorders is the pooled screening yield of 160/47,137 [5].

Diagnostic performance of the components of the examination

Within one screening protocol applied to the same 47,137 persons, the three components of a standard examination performed very differently (Figure 2, Table 2). Electrocardiography interpreted with modern criteria achieved a sensitivity of 94% and specificity of 93% (LR+ 14.8; LR− 0.055), against 20% and 94% for a structured history (LR+ 3.22) and 9% and 97% for physical examination by a specialist (LR+ 2.93) [5]. The false-positive rate was lowest for the instrumental

test (6%) and highest for physical examination (10%) [5]. A clinical examination without the instrumental component therefore misses most of the target condition while generating more false alarms, and a negative result from history or inspection alone barely changes the probability of disease (LR− 0.85 and 0.93) [21, 22].

Table 2. Diagnostic performance of the components of a health examination and agreement between the sources of information it relies on

Component of the examination	Source (population)	Performance against the reference standard			Operational consequence
		Sensitivity	Specificity	Likelihood ratio	
A. Detection of a defined target condition					
12-lead electrocardiography, modern criteria	Harmon et al., 2015 (47,137 young persons) [5]	94%	93%	LR+ 14.8; LR− 0.055	False-positive rate 6%; detects 1 lethal disorder per 294 examined
Structured history questionnaire	Harmon et al., 2015 [5]	20%	94%	LR+ 3.22; LR− 0.85	False-positive rate 8%; misses four of five target disorders
Physical examination by a specialist	Harmon et al., 2015 [5]	9%	97%	LR+ 2.93; LR− 0.93	False-positive rate 10%; a negative result changes probability little
B. Agreement between sources of the same information					
Medical record vs patient-completed questionnaire: type 2 diabetes	Carroll et al., 2017 (775 primary-care patients)	24% vs 36% recorded	κ 0.47	—	Moderate agreement; record alone understates family risk
… colorectal cancer	Carroll et al., 2017	12% vs 14% recorded	κ 0.51	OR for screening 1.9 (1.1–3.1)	p=0.015†; documented risk drives subsequent testing
… coronary artery disease; breast cancer	Carroll et al., 2017	35% vs 22%; 21% vs 22%	κ 0.23; 0.24	—	Fair agreement only; the two sources are not interchangeable
Self-report vs electronic record of exacerbation events	Abrham et al., 2024 (487 adults) [7]	42% of record-identified events unreported	κ 0.47±0.07	—	p<0.001; 33% of self-reported events absent from the record

† Calculated by the present authors. κ, Cohen's kappa (≤0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate); LR+, positive likelihood ratio; LR−, negative likelihood ratio.

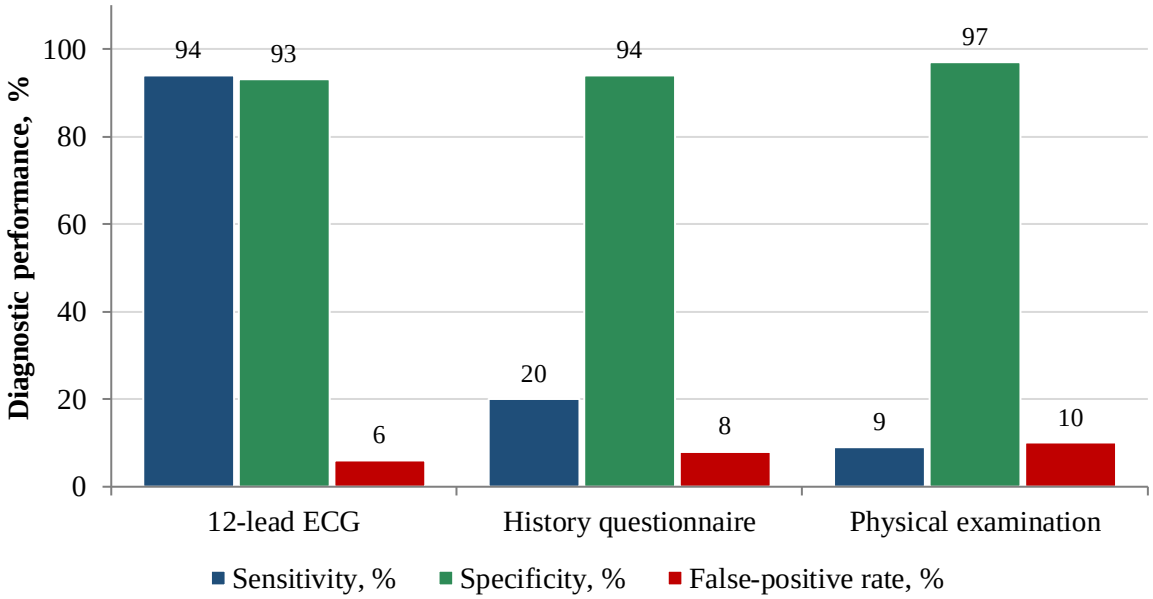


Figure 2. Sensitivity, specificity and false-positive rate of the three components of a cardiovascular screening examination, pooled across 15 studies and 47,137 persons. Data from Harmon et al. [5].

Agreement between the sources of information

Because the commission works from documents as well as from the person in front of it, the agreement between those sources determines what is known at the moment of decision. In 775 primary-care patients, the family history recorded in the electronic record and the history given by the patient agreed only moderately for diabetes (24% vs 36% documented; κ 0.47) and colorectal cancer (12% vs 14%; κ 0.51), and only fairly for coronary artery disease (35% vs 22%; κ 0.23) and breast cancer (21% vs 22%; κ 0.24) (Figure 3). Documentation mattered for subsequent action: a recorded family history predicted colorectal screening (OR 1.9; 1.1–3.1; p=0.015†). The direction of error is not systematic. In 487 adults, 42% of events identified in the electronic record had not been reported by the patient, while 33% of self-reported events were absent from the record (κ 0.47±0.07; p<0.001) [7]. Neither source alone is sufficient, and a prevalence estimate obtained from either is not interchangeable with the other.

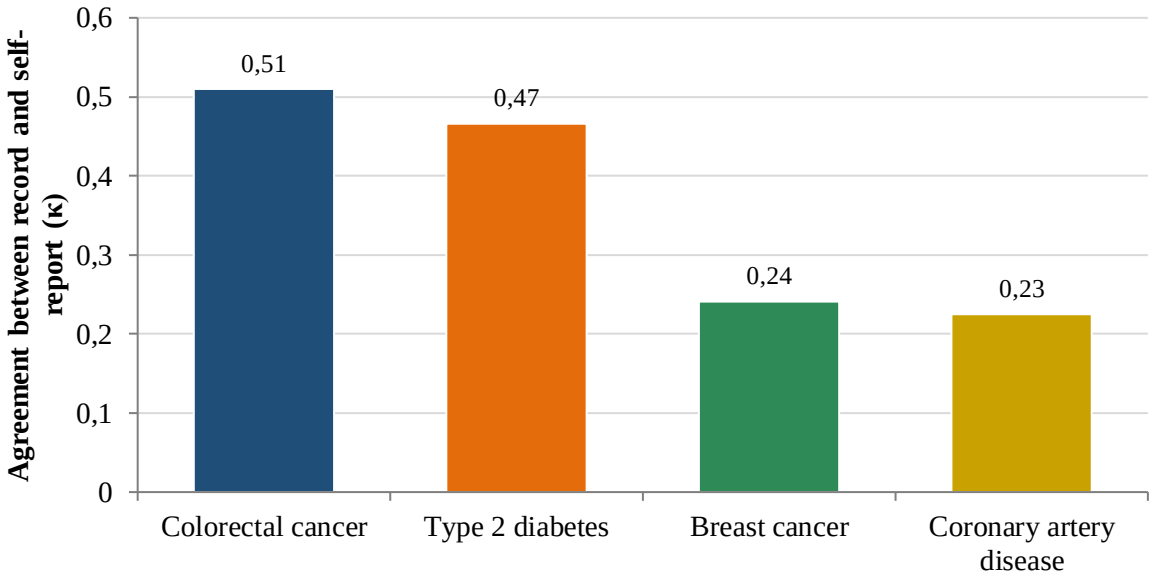


Figure 3. Agreement (Cohen's κ) between the medical record and the patient-completed questionnaire for four conditions, 775 primary-care patients. Values below 0.40 denote fair agreement only. Data from Carroll et al..

Interventions that change performance

The largest randomised evidence base relevant to the management of this process concerns feedback on measured performance (Table 3). In 292 randomised trials with 678 arms, audit and feedback increased the proportion of desired practice by a mean of 6.2 percentage points (95% CI 4.1–8.2; OR 1.47, 1.31–1.64; p<0.001†; moderate certainty), with a median absolute improvement of 2.7% and an interquartile range of 0.0–8.6% [11]; the earlier version of the review reported a median adjusted risk difference of 4.3% (IQR 0.5–16) across 140 studies [23]. Effects were larger when baseline performance was low, when the data described the individual rather than the team, when a local champion with an existing relationship delivered them, and when an explicit action plan was attached [11].

Interventions acting on the physical process produced larger effects on service outcomes: progressive marching distances combined with adapted physical training reduced injury incidence by 33% and attrition by 53% among 651 Swiss recruits [24]. At the system level, twelve years of organised treatment and health-improvement work among adolescents and conscripts in one Russian region accompanied a rise in fitness for service from 63.3% to 77.7%, an absolute gain of 14.4 percentage points (13.6–15.2; OR 2.02, 1.94–2.11; p<0.001)†, although the design is ecological and the conscript cohort shrank by 29.3% over the same period [12]. Screening of a whole cohort at enrolment is operationally feasible but inherits the predictive-value problem of low-prevalence testing [25].

Table 3. Interventions acting on the examination process and their measured effects

Intervention	Source (design)	Evidence base		Effect on the measured outcome	*p*	Certainty / effect modifiers
		Studies	Arms			
Audit and feedback on individual performance	Ivers et al., 2025, Cochrane SR/MA [11]	292	678	Desired practice: +6.2 pp (4.1–8.2); OR 1.47 (1.31–1.64)	<0.001†	Moderate (GRADE). Larger when baseline performance is low, data are individual, a local champion delivers them and an action plan is attached
… earlier version of the same review	Ivers et al., 2012, Cochrane SR/MA [23]	140	108	Median adjusted risk difference +4.3% (IQR 0.5–16)	—	Consistent direction across two decades of trials
Progressive physical load and adapted training	Roos et al., 2015 (controlled field study, Switzerland) [24]	1	4 companies	Injury incidence −33%; attrition −53%	<0.05	651 recruits; binary logistic regression with adjustment for fitness and smoking
Systematic health-improvement work before conscription	Borshchuk et al., 2025 (12-year regional series) [12]	1	2010 vs 2021	Fitness for service 63.3% → 77.7%; +14.4 pp (13.6–15.2); OR 2.02 (1.94–2.11)	<0.001†	Ecological design; concurrent 29.3% fall in the size of the conscript cohort
Pre-enrolment screening with structured interview and quarantine	Norheim et al., 2021 (mass screening, Norway) [25]	1	1,170	2 positive cases at enrolment	—	Illustrates the predictive-value problem of testing in a low-prevalence cohort

† Calculated by the present authors (Altman–Bland method). GRADE, Grading of Recommendations Assessment, Development and Evaluation; IQR, interquartile range; OR, odds ratio; pp, percentage points; SR/MA, systematic review/meta-analysis.

What the current pathway measures

Content analysis of the regulatory corpus identified eight stages with a defined responsible body (Table 4). Five are implemented entirely or predominantly by civilian facilities, and the regulation sets quantitative norms for them: records within ten days, no more than 50 examinations per commission per working day and 100 at the assembly point, a conclusion on the day of examination, free additional examination, and deferment for treatment up to twelve months. Measurement, however, begins only at stage 6. Returns from the assembly point generate an official letter and an internal inquiry, and each serviceman discharged as wrongly conscripted triggers a joint inquiry with disciplinary consequences and a report by 1 June. Of the fifteen candidate indicators that the pathway would support, four are already collected in some form and eleven are not collected at all.

Table 4. The regulatory pathway of conscript medical examination in the Republic of Uzbekistan and the point at which quality is first measured

Stage	Content	Responsible body		Quantitative norm set by the regulation	What is measured
		Civilian healthcare	Defense authority		
1	Annual examinations of males aged 15–16 years, immunisation, health-improvement measures	Territorial health bodies; outpatient facilities	District defense office (registration)	Annual; records and reporting defined jointly by the ministries	Nothing beyond completion
2	Mandatory diagnostic tests, oral sanitation, transfer of medical records	Outpatient facilities	District defense office (requests)	Records provided within 10 days	Nothing
3	Examination at registration and conscription; fitness categories A–F	Specialists of eight profiles in the commission	Territorial medical commission	≤50 persons per working day; conclusion on the day of examination	Category issued
4–5	Additional examination and treatment before conscription	Outpatient facilities and hospitals, free of charge	Referral by the commission	Deferment for treatment up to 12 months	Nothing beyond repeat examination
6	In-depth control examination before dispatch to the troops	—	Military medical commission of the assembly point	≤100 persons per day	Returns → official letter and internal inquiry
7–8	Transfer of the insert sheet to the unit; discharge of wrongly conscripted servicemen	Territorial health bodies (joint inquiry)	Military medical commissions; central commission	Inquiry materials submitted by 1 June annually	Wrongful conscription → disciplinary measures

Compiled from Presidential Decree PD-4076 of 25 December 2018 (as amended) and Presidential Decree PD-4630 of 6 March 2020.

Discussion

Principal findings

Three findings follow from this synthesis. First, the population examined is not healthy, and the most frequent conditions in it — refractive, musculoskeletal, metabolic and psychiatric — are precisely those that a one-day inspection detects poorly [1–4]. Second, within a single protocol the instrumental component of an examination outperforms its clinical components by a factor of five to ten in sensitivity [5], while the documentary and self-reported sources of the same information agree at κ 0.23–0.51 [7]; measured quality therefore depends on the method of ascertainment before it depends on the diligence of the examiner. Third, the one intervention with a large randomised evidence base — feedback on individual measured performance — yields a modest but consistent improvement (+6.2 pp) that is greatest exactly where baseline performance is lowest [11], which is the situation of an unmeasured process [26].

A process that is governed but not measured

The Uzbek pathway is complete in the legal sense and incomplete in the Donabedian sense [10]: it assigns responsibility for every stage, sets deadlines and caps, and guarantees free additional outpatient and inpatient examination, yet it records no indicator of structure or process. The two outcome measures that exist — returns and wrongful conscription — arise after the civilian stages are over and are attached to sanctions. Feedback of that kind is poorly designed for improvement: the randomised evidence shows that effects depend on individual-level data delivered by a trusted local source with an action plan, not on retrospective attribution of blame [11]. The same evidence predicts that introducing process indicators in a system that currently has none should produce more than the pooled average, because baseline performance is low [11, 23, 27, 28].

Ascertainment before interpretation

The agreement data (Figure 3) have a direct operational reading. If a commission relies on a conscript’s account where the record is silent, or on the record where the conscript reports more, it will produce different fitness categories for the same person, and the resulting discordance between the territorial commission and the assembly-point commission will be interpreted as an error by one of them. This is not primarily a disciplinary problem but a measurement problem, and it is bounded by the reliability of the sources, not by the competence of the examiner. The corollary for indicator design is that any indicator describing detection must specify the instrument, and that instrumentally verified findings should take precedence over self-reported ones wherever a device exists — as the ECG comparison shows most sharply [5].

An operational indicator framework

Combining the four domains yields the framework in Table 5: fifteen indicators, of which three describe structure, six process and six outcome, each tied to a source of data that the pathway already generates and to a managerial action that could change it. Four indicators (10–13) can be computed from documents that are currently filed only for inquiry purposes; nine require no new instrument beyond what the Regulation already mandates. The framework is a hypothesis, not a standard: target values must be set empirically, and the indicators must be shown to be feasible, reliable and sensitive to change before they are used for comparison between facilities.

Table 5. Operational indicator framework for the quality of medical examination of conscription-age men (hypothesis for prospective testing)

No.	Component	Indicator	Measurement		Evidence that it is actionable
			Unit and source	Frequency and level	
A. Structure					
1	Structure	Staffing with the eight specialist profiles taking part in examinations	%; staffing schedule	Annually; facility	Daily cap of 50 examinations presupposes a complete panel
2	Structure	Capacity for mandatory tests, including 12-lead electrocardiography	%; equipment inventory	Annually; facility	ECG sensitivity 94% vs 9–20% for clinical components [5]
3	Structure	Electronic record with a data channel to the defense office	Yes/no; information system	Annually; facility	Record–self-report agreement only κ 0.23–0.51 [7]
B. Process					
4	Process	Coverage of males aged 15–16 years by annual examinations	%; facility reports	Quarterly; facility, district	Organised pre-conscription work preceded +14.4 pp in fitness [12]
5	Process	Completeness of mandatory diagnostic tests before examination	%; examination cards	Quarterly; facility	Incomplete testing removes the only high-sensitivity component [5]
6	Process	Records supplied to the defense office within the 10-day limit	%; request register	Quarterly; facility	Regulatory norm, currently unmeasured
7	Process	Conscripts referred for additional examination	%; commission protocols	Per campaign; district	Marks the boundary between screening and diagnosis
8	Process	Treatment completed before the start of the campaign	%; treatment register	Per campaign; facility	Deferment is bounded at 12 months
9	Process	Interval from detection of disease to completion of treatment	Days; outpatient cards	Per campaign; facility	Half of service back-pain episodes arise in the first months [2]

No.	Component	Indicator	Measurement		Evidence that it is actionable
			Unit and source	Frequency and level	
C. Outcome					
10	Outcome	Conditions first detected at the commission rather than in primary care	%; protocols and cards	Per campaign; district	Direct measure of the detection gap [12]
11	Outcome	Discordance between territorial and assembly-point commissions	%; assembly-point reports	Per campaign; district	Agreement between two assessments of one person
12	Outcome	Conscripts returned from the assembly point for health reasons	%; official letters	Per campaign; district	Already collected, currently used only for sanctions
13	Outcome	Servicemen discharged as wrongly conscripted	Per 1,000; inquiry files	Annually; district	Already collected
14	Outcome	Fitness for service, categories A and B	%; commission reports	Per campaign; district	Modifiable: 63.3% → 77.7% over 12 years [12]
15	Outcome	Attrition and medical discharge during the first six months of service	%; unit medical service	Per cohort; region	Attrition 10% with identifiable predictors [18]; reducible by 53% [24]

Target values are to be established empirically before the indicators are used for comparison between facilities. Indicators 10–13 can be computed from documents the pathway already generates. ECG, electrocardiography.

Implications for study design

These findings inform a prospective evaluation in the Republic of Uzbekistan in which outpatient facilities of districts are allocated to indicator-based audit and feedback or to current practice, with coverage, completeness of mandatory testing and the proportion of conditions first detected at the commission as co-primary indicators and a 12-month follow-up of returns and medical discharge. Using the pooled audit-and-feedback effect as the design assumption (6.2 pp; 95% CI 4.1–8.2) [11] and allowing for the low baseline that increases it, a cluster design with facilities per arm and an intraclass correlation of 0.05 would detect an absolute difference of 10 percentage points with 80% power at two-sided $\alpha=0.05$. Conscript health data collected at a single legally fixed moment also constitute a population sampling frame of a quality rarely available in young men [6, 19], and the same instrument can serve surveillance as well as selection.

Limitations

This is a structured rather than a full systematic review: PubMed/MEDLINE was the principal database, extraction was performed by one reviewer with re-checking, the protocol was not registered, and risk of bias was not reassessed beyond the source ratings. The evidence on diagnostic performance and on agreement comes from populations outside the conscription setting and from a cardiovascular and primary-care context; its transfer to a general fitness examination is an assumption, not a finding. The regional fitness series is ecological and confounded by a concurrent fall in cohort size [12]. Several derived *p* values and odds ratios rest on counts reconstructed from rounded percentages and are approximate. No empirical data from the Republic of Uzbekistan are presented; the framework is therefore untested in the setting for which it is proposed.

Conclusion

Medical examination of conscription-age men in the Republic of Uzbekistan is a clinical process delivered by primary care under legal deadlines, but it is managed as an administrative procedure: its outcomes are recorded only after the civilian stages are complete and are used for attribution of fault. The population examined carries a substantial and partly invisible burden of disease, the components of the examination differ by an order of magnitude in sensitivity, and the sources of information on which commissions rely agree only fairly to moderately. Feedback on measured individual performance improves professional practice modestly but reliably, and most where measurement is currently absent. A fifteen-indicator structure–process–outcome framework built on data the pathway already generates, with instrumentally verified inputs taking precedence over self-reported ones, is therefore a justified but untested hypothesis; prospective cluster-randomised evaluation is the necessary next step.

Additional information

Conflict of interest. The authors declare no conflict of interest.

Ethical approval. Not required: the review used published aggregate data and public regulatory documents only.

Author contributions. All authors contributed to the compilation and layout of this material.

Data availability. All data are derived from the published sources cited; the calculations underlying derived statistics are available from the corresponding author on request.

References

1. Pärssinen O., Franssila M.L., Nordhausen K., et al. The prevalence of myopia in Finnish conscripts has not increased in recent decades and is lower than in many East and Southeast Asian countries. Invest Ophthalmol Vis Sci. 2025;66(12):10. https://doi.org/10.1167/iovs.66.12.10
2. Suikkanen S., Pihlajamäki H., Parviainen M., Kautiainen H., Kiviranta I. Prevalence of and risk factors for back pain among young male conscripts during compulsory Finnish military service. Mil Med. 2023;188(3–4):e739–e744. https://doi.org/10.1093/milmed/usab375
3. Yang L., Juan A., Waldhoer T. Prevalence and trends in obesity among Austrian conscripts from 1983 to 2017. Wien Klin Wochenschr. 2021;135(13–14):358–363. https://doi.org/10.1007/s00508-021-01941-9
4. Yeom C.W., Oh G.H., Jung S., et al. Prevalence and comorbidities of adult ADHD in male military conscripts in Korea. Psychiatry Res. 2020;293:113401. https://doi.org/10.1016/j.psychres.2020.113401
5. Harmon K.G., Zigman M., Drezner J.A. The effectiveness of screening history, physical exam, and ECG to detect potentially lethal cardiac disorders in athletes: a systematic review/meta-analysis. J Electrocardiol. 2015;48(3):329–338. https://doi.org/10.1016/j.jelectrocard.2015.02.001
6. Sperhacke R.D., da Motta L.R., Kato S.K., et al. HIV prevalence and sexual behavior among young male conscripts in the Brazilian army, 2016. Medicine (Baltimore). 2018;97(1S Suppl 1):S25–S31. https://doi.org/10.1097/MD.0000000000009014
7. Abrham Y., Zeng S., Lin W., et al. Self-report underestimates the frequency of the acute respiratory exacerbations of COPD but is associated with BAL neutrophilia and lymphocytosis: an observational study. BMC Pulm Med. 2024;24(1):433. https://doi.org/10.1186/s12890-024-03239-8
8. Holden B.A., Fricke T.R., Wilson D.A., Jong M., Naidoo K.S., Sankaridurg P., et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*. 2016;123(5):1036–1042. https://doi.org/10.1016/j.ophtha.2016.01.006
9. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet*. 2017;390(10113):2627–2642. https://doi.org/10.1016/S0140-6736(17)32129-3

10. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260(12):1743–1748. https://doi.org/10.1001/jama.1988.03410120089033
11. Ivers N., Yogasingam S., Lacroix M., et al. Audit and feedback: effects on professional practice. Cochrane Database Syst Rev. 2025;3(3):CD000259. https://doi.org/10.1002/14651858.CD000259.pub4
12. Borshchuk E.L., Kuzmin S.A., Grigorieva L.K., Izbagambetova K.A., Duisembaeva A.N., Vinogradova M.T. Assessment of the state of health, lifestyle and quality of life of citizens subject to military conscription. Zdorov'e Naseleniya i Sreda Obitaniya. 2025;33(2):21–29. https://doi.org/10.35627/2219-5238/2025-33-2-21-29 (In Russ.).
13. Chernov D.A. Assessment of the health status of conscripts. Problemy Zdorov'ya i Ekologii. 2021;18(1):5–13. https://doi.org/10.51523/2708-6011.2021-18-1-1 (In Russ.).
14. Belyakova E.V., Liferov R.A. Socio-hygienic features of health of adolescents of pre-conscription and conscription age. Nauka Molodykh (Eruditio Juvenium). 2024;12(1):135–143. https://doi.org/10.23888/HMJ2024121135-143 (In Russ.).
15. Page M.J., McKenzie J.E., Bossuyt P.M., et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71
16. Altman D.G., Bland J.M. How to obtain the P value from a confidence interval. BMJ. 2011;343:d2304. https://doi.org/10.1136/bmj.d2304
17. Carrasco A., Navren M., Larsson I., Taube F., Björkman F. Systematic review and meta-analysis: iron deficiency and iron deficiency anemia among military recruits undergoing basic combat training. Mil Med. 2025;190(3–4):541–553. https://doi.org/10.1093/milmed/usae437
18. Gendler S., Talmy T., Shapiro M., et al. Predictors of combat training attrition in Israel Defense Forces soldiers. Occup Med (Lond). 2023;73(2):80–84. https://doi.org/10.1093/occmed/kqad008
19. Crump C., Sundquist J., Winkleby M.A., Sieh W., Sundquist K. Physical fitness among Swedish military conscripts and long-term risk for type 2 diabetes mellitus: a cohort study. Ann Intern Med. 2016;164(9):577–584. https://doi.org/10.7326/M15-2002
20. Knapik J.J., Sharp M.A., Canham-Chervak M., Hauret K., Patton J.F., Jones B.H. Risk factors for training-related injuries among men and women in basic combat training. *Med Sci Sports Exerc*. 2001;33(6):946–954. https://doi.org/10.1097/00005768-200106000-00014
21. Drezner J.A., Sharma S., Baggish A., Papadakis M., Wilson M.G., Prutkin J.M., et al. International criteria for electrocardiographic interpretation in athletes: consensus statement. *Br J Sports Med.* 2017;51(9):704–731. https://doi.org/10.1136/bjsports-2016-097331
22. Corrado D., Basso C., Pavei A., Michieli P., Schiavon M., Thiene G. Trends in sudden cardiovascular death in young competitive athletes after implementation of a preparticipation screening program. *JAMA*. 2006;296(13):1593–1601. https://doi.org/10.1001/jama.296.13.1593
23. Ivers N., Jamtvedt G., Flottorp S., et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012;6:CD000259. https://doi.org/10.1002/14651858.CD000259.pub3

24. Roos L., Boesch M., Sefidan S., Frey F., Mäder U., Annen H., et al. Adapted marching distances and physical training decrease recruits' injuries and attrition. Mil Med. 2015;180(3):329–336. https://doi.org/10.7205/MILMED-D-14-00184

25. Norheim A.J., Borud E.K., Lind A., et al. Pre-screening and preventive quarantine likely explains the low SARS-CoV-2 prevalence among Norwegian conscripts. Scand J Prim Health Care. 2021;39(1):31–34. https://doi.org/10.1080/02813432.2021.1880101

26. Kringos D.S., Boerma W.G., Hutchinson A., van der Zee J., Groenewegen P.P. The breadth of primary care: a systematic literature review of its core dimensions. *BMC Health Serv Res*. 2010;10:65. https://doi.org/10.1186/1472-6963-10-65

27. Campbell S.M., Braspenning J., Hutchinson A., Marshall M.N. Research methods used in developing and applying quality indicators in primary care. *BMJ*. 2003;326(7393):816–819. https://doi.org/10.1136/bmj.326.7393.816

28. Mainz J. Defining and classifying clinical indicators for quality improvement. *Int J Qual Health Care*. 2003;15(6):523–530. https://doi.org/10.1093/intqhc/mzg081

Information about the authors

Shoira Q. Abdusattarova

Saidakhon Sh. Mirrakhimova