RRL-QMS Working Paper (JRS) — v1.0

A Multi-Level Regulatory Readiness Framework (RRL-QMS) and Predictive Modelling for Quality Maturity in Medical Device Approvals

Keywords: Medical Devices; Regulatory Readiness; RRL-QMS; Quality Maturity Model; Approval Prediction; ISO 13485

Abstract

Background: Global regulators now demand auditable Quality Management Systems (QMS) for medical devices, yet start-ups lack a metric that links QMS maturity to approval probability. Objective: We present the nine-level Regulatory Readiness Level-QMS (RRL-QMS) framework and a predictive model that quantifies how maturity, risk class and prior experience shape first-pass approval. Methods: RRL-QMS was mapped to ISO 13485:2016 and MDSAP audit tasks. A simulated cohort of 500 projects (RRL-1-RRL-9) plus ten real, anonymised cases was analysed. Logistic regression with cross-validation, calibration (Brier score) and ROC metrics assessed performance; sensitivity and decision-curve analyses examined robustness and utility. Results: Each single-level rise in RRL-QMS almost doubled approval odds (OR 1.72, $p < 0.001$). The model showed strong discrimination (AUC 0.83) and good calibration (Brier 0.166). RRL-QMS explained ~88 % of predictive variance and provided stable predictions for 84 % of projects; forecasts matched outcomes in nine of ten real cases. Conclusion: RRL-QMS couples a staged, standards-aligned QMS ladder with a validated prediction engine, enabling developers and regulators to benchmark readiness, prioritise gaps and make risk-based submission decisions.

Introduction

Global regulatory expectations for medical devices have escalated in recent years [20]. New regulations like the EU Medical Device Regulation (MDR 2017/745) (and its counterpart for diagnostics, IVDR 2017/746) and stricter enforcement of the U.S. FDA's Quality System Regulation (21 CFR Part 820) are raising the compliance burden on manufacturers. International audit schemes such as the Medical Device Single Audit Program (MDSAP) now serve as unifying frameworks across multiple jurisdictions (USA, Canada, Australia, Brazil, Japan). Early-stage device companies, however, often lack tools to measure or communicate their regulatory preparedness. Currently, no standardized maturity model quantitatively links the development of a QMS with the likelihood of regulatory approval. Existing models like Technology Readiness Levels (TRLs) or the Capability Maturity Model Integration (CMMI) focus on technological or process maturity but do not capture regulatory progress. They also provide no way to estimate approval success from a given maturity profile. Even recently introduced quality maturity initiatives - such as the FDA's Quality Management Maturity (QMM) program and the Medical Device Innovation Consortium's Medical Device Discovery Appraisal Program (MDDAP) - lack a multi-level predictive framework tied explicitly to submission readiness. This gap in readiness assessment can lead to missed timelines, failed submissions, or extensive remediation efforts after audits.

To address this gap, we propose a comprehensive Regulatory Readiness Level-QMS (RRL-QMS) framework. The RRL-QMS defines an ascending ladder of QMS maturity divided into nine levels aligned with international standards-most notably ISO 13485:2016-and with the Medical Device Single Audit Program (MDSAP) audit model. Each level corresponds to concrete QMS milestones and documentary evidence of compliance. In tandem, we develop a predictive statistical model that quantifies how an organisation's RRL-QMS level, together with device risk class and prior regulatory experience, affects its probability of regulatory approval.

Research Question (RQ). To what extent does organisational QMS maturity-as operationalised by the nine-level RRL-QMS ladder-predict the probability of first-pass regulatory approval for medical devices?

Accordingly, this study develops and validates a calibrated logistic-regression model that quantifies that relationship.

This study makes four main contributions:

Framework Development: We formulate a structured nine-level RRL-QMS ladder with explicit criteria mapped to the MDSAP audit domains, providing a staged roadmap from regulatory awareness to post-market excellence.

Predictive Model: We build and validate a model that links an organisation's RRL-QMS level, device risk class and prior experience to its approval likelihood, thereby turning qualitative maturity assessment into quantitative decision support.

Rigorous Evaluation: We demonstrate strong out-of-sample performance (AUC ≈ 0.83) and robust calibration through k-fold cross-validation, Brier score, decision-curve and global sensitivity analyses, complemented by a 10-case real-world pilot.

Comparative Analysis: We position RRL-QMS vis-à-vis established frameworks (TRL, CMMI, FDA-QMM, MDDAP, ISO 13485) and show how it uniquely integrates staged regulatory focus with predictive analytics.

By bridging the gap between qualitative compliance check-lists and quantitative outcome prediction, RRL-QMS offers a novel evidence-based tool for industry and regulators to gauge "regulatory readiness" in a standardised manner.

Methods

RRL-QMS Framework Construction

We first conducted a broad review of international QMS requirements to inform the design [3, 4] of the RRL-QMS framework. Key regulations, standards, and guidance documents from around the world (covering at least 10 jurisdictions across North America, Europe, and Asia-Pacific) were analyzed to extract common compliance domains. Universal foundations such as achieving ISO 13485:2016 certification and implementing risk management per ISO 14971 were included, alongside region-specific requirements (for example, Japan's additional QMS ordinance and China's YY/T 0287 standard, which mirrors ISO 13485). From this analysis, we defined ten principal QMS domains encompassing management responsibility, design and design transfer controls, risk management, document control, supplier management, production and service controls, corrective and preventive actions (CAPA), internal auditing, post-market surveillance, and others.

Using these domains, we developed a self-assessment checklist of 50 items to measure compliance in each area. Each checklist item is a binary (yes/no) question reflecting a specific regulatory requirement (e.g., "Is there a documented risk management procedure compliant with ISO 14971?"). For each organization assessed, domain scores are calculated as the fraction of items met in each domain, and an overall compliance percentage is obtained via a weighted sum of domain scores (with higher weight given to more critical domains). This overall score is then mapped to a Regulatory Readiness Level (RRL) on an ordinal scale representing increasing maturity. In our framework, RRL-1 corresponds to an initial or ad-hoc QMS (very low compliance), while RRL-9 represents a fully realized, highly mature QMS that embodies continuous improvement and meets global regulatory expectations. Each intermediate level is defined by concrete criteria and deliverables that should be in place before advancing. For example, reaching RRL-4 requires having a full design control process established (with at least one completed design review), and RRL-7 signifies that the organization has passed an external ISO 13485/MDSAP audit and obtained regulatory approval for a product. These criteria ensure that each step on the RRL-QMS ladder reflects a meaningful increase in regulatory capability.

Data Simulation

Because large real-world datasets of start-ups with varied QMS maturity and known regulatory outcomes are not readily available, we created a synthetic dataset for the predictive modeling component. We simulated N = 500 hypothetical medical device development projects. Each simulated project was assigned: (i) an RRL-QMS maturity level (1 through 9) representing its QMS state at the time of regulatory submission; (ii) a device risk class (for example, comparing a low-risk Class I device vs. a high-risk Class III device); (iii) a binary flag for prior regulatory experience (indicating whether the development team had at least one prior product approval); and (iv) an outcome of either regulatory approval (“pass”) or failure (e.g. rejection or major deficiency letter). The simulation was designed such that higher RRL levels and other favorable factors conferred a higher probability of approval. Specifically, we assumed that strong performance in critical QMS domains (such as management responsibility and risk management) would significantly increase the chance of success. We generated an underlying “true” model for the approval probability as a logistic function of the various inputs (with larger weights for RRL and certain domains), then sampled binary outcomes with randomness to mimic real-world uncertainty. This process resulted in approximately 65% of simulated projects being approvals and 35% failures, which is in line with realistic success rates for well-prepared organizations. To ensure robustness of our findings, we repeated the entire simulation procedure 10 times with different random seeds (producing 10 independent datasets) and later verified that our results were consistent across these replicates.

Predictive Model Development

We fit a logistic regression model to predict the probability of regulatory approval for a given project based on its RRL-QMS level, device risk class, and prior regulatory experience. Logistic regression was chosen for its interpretability and suitability for binary (pass/fail) outcomes. The model takes the form:

logP(approval)1−P(approval)=β0+β1⋅(RRL level)+β2⋅(Risk class)+β3⋅(Prior experience),\log \frac{P(\text{approval})}{1 - P(\text{approval})} = \beta_0 + \beta_1 \cdot (\text{RRL level}) + \beta_2 \cdot (\text{Risk class}) + \beta_3 \cdot (\text{Prior experience}),

where the coefficients $\beta_1, \beta_2, \beta_3$ capture the influence of each predictor. Here, Risk class was coded as an indicator variable (e.g. 1 for a high-risk device and 0 for a low-risk device), and Prior experience was binary (1 if the team had prior approval experience, 0 if not). We allocated 80% of the 500 simulated cases for model training and reserved 20% as an independent test set. Model fitting was performed using maximum likelihood estimation. To prevent overfitting, we applied a mild L2 regularization (ridge penalty) to the logistic model, with the regularization strength tuned via 5-fold cross-validation on the training data. This procedure guards against unstable coefficient estimates in the presence of correlated inputs. The final model’s learned coefficients aligned with expectations: for instance, the RRL-QMS level received a substantial positive weight (indicating higher readiness strongly boosts approval odds), while a high device risk class had a negative weight (indicating that high-risk devices are inherently harder to approve), reflecting the intuitive direction of each factor’s effect.

Model Evaluation and Validation

After training, we evaluated the logistic model on the held-out test set to assess its predictive performance. We computed the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) to measure discriminative ability. The model achieved a high AUC, indicating strong ability to distinguish between approvals and failures (an AUC of 0.5 would be no better than chance). We also assessed calibration by comparing predicted approval probabilities against actual outcome frequencies: a calibration plot was constructed, and the Brier score (mean squared error of the probability predictions) was calculated as an overall measure of calibration quality. Additionally, we performed Decision Curve Analysis to gauge the practical utility of the model for decision-making. This analysis involves comparing the net benefit of using the model to guide decisions (e.g. whether to invest in additional QMS improvements or to proceed with submission) against default strategies of intervening in all cases or in none. We examined a range of threshold probabilities (the minimum predicted risk of failure at which one would take preventive action) to determine where the model adds positive net benefit.

Beyond these standard metrics, we carried out a Sobol global sensitivity analysis [7] on the fitted model to understand the contribution of each input factor to output variability. By varying each input (RRL level, risk class, prior experience) across its range and averaging over the distributions of the others, we estimated Sobol sensitivity indices - essentially quantifying the percentage of the model's output variance attributable to each factor. This helps identify the dominant drivers of approval outcomes. We also examined model stability by comparing results across the 10 different simulated datasets (the independent simulation replicates mentioned above). We define a Stability Index as the proportion of cases for which the predicted approval probability does not vary excessively across the different model iterations. In practice, we found that for a large majority of projects, the predictions were very consistent between runs. Finally, we performed an external validation by applying the model to a small pilot dataset of 10 real-world medical device projects (drawn from anonymized case studies with known regulatory outcomes). These cases were not used in any part of model development. We recorded the model's predicted approval probabilities for each and checked them against the actual outcomes to see if high predictions corresponded to actual approvals and low predictions to failures. This served as an initial real-world check on the model's applicability.

All data simulation and analysis code was developed in Python. In line with open science principles, we ensured the analysis workflows are reproducible. However, the proprietary 50-item checklist content, the full codebase, and detailed real-world data are not disclosed in this publication; they remain available for review under a formal licensing agreement, as noted above.

Results

RRL-QMS Framework Overview

The RRL-QMS framework is a nine-step "ladder" that maps the evolution of a medical-device Quality Management System (QMS) from first regulatory awareness to post-market excellence. Each level has clear entry/exit criteria, deliverables, and documentary evidence aligned with ISO 13485:2016 clauses and MDSAP audit tasks. Progression is strictly sequential: an organization must satisfy every requirement at a given level before it can claim the next. A concise description of all nine levels follows.

Table 1. RRL-QMS Maturity Levels and Key Achievements

Level	Name	Key Achievements
RRL-1	Regulatory Awareness	Company has identified the intended use of its device and is aware of basic regulatory obligations, but no formal QMS artefacts exist.
RRL-2	Foundational QMS Drafted	Quality Policy and draft Quality Manual are approved; document-control and risk-management procedures are written; key quality roles (e.g., Management Rep) are assigned; introductory regulatory training records are filed; initial controlled■document list created.
RRL-3	SOP Implementation	Core SOPs (document control, design control, training) are issued and in routine use; basic design-history records and change-control logs exist; risk-management plan is active; first internal training and document-control cycle completed.

Level	Name	Key Achievements
RRL-4	Design Controls Operational	Full design-control process is executed at least once; DHF opened and one structured design review held; traceability matrix drafted; preliminary supplier-quality files started; early CAPA log initiated.
RRL-5	Production-Quality Practices	Supplier qualification/agreements in place; CAPA system has closed at least one corrective action; production and service controls documented; device master record (DMR) compiled; first lot/batch records demonstrate GMP traceability.
RRL-6	Pre-Certification / Submission-Ready	All QMS processes have run a complete cycle; critical production and software systems are validated (IQ/OQ/PQ); organisation-wide internal audit and management review completed with major findings closed; full technical file (EU MDR Annex II) or 510(k)/PMA draft compiled and gap-assessed; external certification audit (ISO 13485 or MDSAP Stage 1) scheduled.
RRL-7	Certified & First Approval	Organisation has passed an external ISO 13485 or MDSAP audit and obtained at least one market authorisation (e.g., CE-marking, 510(k) clearance); post-approval surveillance plan activated.
RRL-8	Global Compliance & Post-Market Excellence	Multi-region licences maintained; post-market surveillance system producing trend reports; periodic re-certification audits passed; feedback loops from field data drive design and process updates.
RRL-9	Continuous Improvement	Quality is embedded company-wide; KPI-driven optimisation of every QMS domain; advanced analytics (e.g., SPC, predictive CAPA) in routine use; organisation routinely scales QMS to new products/markets with minimal non-conformities.

Figure 1. Nine-level RRL-QMS ladder-from RRL-1 “Regulatory Awareness” to RRL-9 “Continuous-Improvement Culture.” Each step marks a key maturity milestone for medical-device quality systems.

Predictive Model Results

The logistic regression model confirmed that higher QMS maturity (as quantified by RRL-QMS level) is strongly associated with regulatory success. Holding other factors constant, each one-level increase in an organization’s RRL-QMS corresponded to approximately a 1.72× increase in the odds of approval (odds ratio ≈ 1.72, $\beta \approx +0.540$ per level, $p < 0.001$). In practical terms, moving from, say, RRL-3 to RRL-4 or from RRL-7 to RRL-8 markedly raises the chances of obtaining approval. Device risk level and team experience also showed significant effects in the model: high-risk devices (e.g. Class III versus

Class I) had about 50% lower odds of approval (OR ≈ 0.51, $\beta \approx -0.670$, $p \approx 0.001$), reflecting the tougher regulatory scrutiny on higher-risk products, while organizations with prior regulatory experience were over 3.5 times more likely to succeed again than first-timers (OR ≈ 3.52, $\beta \approx +1.258$, $p < 0.001$). All three predictors - RRL level, device risk, and prior experience - were statistically significant in the model, underscoring that both quality system maturity and contextual factors independently influence regulatory outcomes.

Overall model performance was encouraging. On the hold-out test dataset, the model achieved an ROC AUC (Area Under the Curve) of ~0.83, indicating a strong ability to discriminate between approvals and failures (for comparison, an AUC of 0.5 denotes no discriminative power). The calibration of the model was also good - the predicted approval probabilities closely matched the observed success rates across different probability bands. The overall Brier score was 0.166, indicating a reasonably low prediction error (a Brier score of 0 would indicate perfect calibration). In a decision curve analysis, the RRL-QMS model provided positive net benefit across a range of risk threshold preferences. In other words, using the model's output to guide decisions (such as whether to delay a submission for further QMS improvements) would yield better expected outcomes than either intervening in all cases or in none, for a broad range of decision-maker risk tolerances. Specifically, the model was most beneficial when the threshold for taking action was around a 60-75% predicted risk of failure: above these thresholds (i.e. when the model indicates a relatively high risk of non-approval), acting on the model's warning (for example, by strengthening the QMS before submitting) clearly outperformed treating all submissions the same.

The model also proved to be robust and stable. When we repeated the entire simulation and modeling process 10 times (each with fresh random data), the results were highly consistent: for approximately 84% of the projects, the predicted approval probability varied only modestly between runs, corresponding to a Stability Index > 0.9 in those cases. This means that the vast majority of predictions were not sensitive to the particular simulation instance, instilling confidence that the model's insights are not an artifact of any one synthetic dataset. The Sobol global sensitivity analysis [7] revealed that the RRL-QMS level alone accounted for roughly 88% of the variance in the model's predicted outcomes. In other words, uncertainty or changes in the RRL-QMS input drive the bulk of the uncertainty in the approval prediction, far more than the other factors - highlighting RRL level as by far the dominant predictor of success. Finally, when we applied the model to the real-world pilot set of 10 medical device projects (anonymized startup cases with known outcomes), the model's predictions aligned with the actual regulatory outcomes in 9 out of 10 cases. In each of those nine cases, if the model predicted a high probability of approval, the device was indeed approved, and if the model predicted a low probability, the submission ultimately failed. This high agreement in a small external test suggests that the RRL-QMS model, despite being trained on synthetic data, captures signals that are pertinent to real regulatory scenarios.

Asia-Pacific Regulatory Annex

Table 2. Comparative QMS Requirements by Jurisdiction

Country	Primary Regulatory Body	Baseline QMS Standard	Notable Additional Requirements	Status vis-à-vis MDSAP*
China	NMPA	ISO 13485 + CN GMP	Mandatory on-site QMS audit; hygiene, layout, staffing criteria	No - local audit only
Japan	PMDA	Ministerial Ordinance 169 (≈ ISO 13485)	Accepts MDSAP certificate with document review	Yes - recognised
South Korea	MFDS	KGMP (ISO-derived)	Mandatory Korean audit; MDSAP not accepted	Observer

Country	Primary Regulatory Body	Baseline QMS Standard	Notable Additional Requirements	Status vis-à-vis MDSAP*
Singapore	HSA	ISO 13485 / MDSAP	Accepts either; SAC accreditation required for ISO	Yes - accepted
Australia	TGA	ISO 13485 / MDSAP	CE certificate or MDSAP report usable; Essential-Principles mapping required	Yes - accepted
Canada	Health Canada	MDSAP certificate mandatory	Applies to Class II-IV devices; ISO alone insufficient	Mandatory

Table 2. Abbreviations - NMPA: National Medical Products Administration; PMDA: Pharmaceuticals and Medical Devices Agency; MFDS: Ministry of Food and Drug Safety; HSA: Health Sciences Authority; TGA: Therapeutic Goods Administration.

Interpretation. The annex highlights heterogeneity across Asia-Pacific: while Japan, Singapore, Australia and Canada accept or require MDSAP certificates, China and South Korea still rely on sovereign GMP audits. These variations justify the inclusion of regional overlays within the RRL-QMS checklist and underscore the framework's utility for firms seeking multi-jurisdictional clearance.

Discussion

Comparison with Existing Frameworks: To put the RRL-QMS approach in context, we compared its characteristics to several well-known technology and quality maturity frameworks. Table 3 provides a side-by-side overview of RRL-QMS versus TRL, CMMI, the FDA's Quality Management Maturity (QMM) program, the Medical Device Discovery Appraisal Program (MDDAP), and ISO 13485. The comparison spans the focus/scope of each framework, their domain of application, structural levels, assessment methods, primary outputs, and typical usage contexts.

Table 3. Comparative characteristics of RRL-QMS vs. other maturity/readiness frameworks

Aspect	RRL-QMS (Regulatory Readiness Level - QMS)	TRL (Technology Readiness Levels)	CMMI (Capability Maturity Model Integration)	FDA-QMM (Quality Management Maturity)	MDDAP (MDIC Medical Device Discovery Appraisal Program)	ISO 13485 (QMS Standard)
Focus/Scope	Regulatory QMS readiness tied to regulatory approval requirements.	Technological development progress of a product (from concept prototype to deployment).	Organizational process capability and maturity (general process quality improvement).	Organizational quality culture and maturity (manufacturing focus in pharma).	Medical device quality system and process maturity (CMMI-based appraisal; FDA/MDIC initiative).	Quality management system compliance for regulatory purposes (baseline regulatory requirements).

Aspect	RRL-QMS (Regulatory Readiness Level - QMS)	TRL (Technology Readiness Levels)	CMMI (Capability Maturity Model Integration)	FDA-QMM (Quality Management Maturity)	MDDAP (MDIC Medical Device Discovery Appraisal Program)	ISO 13485 (QMS Standard)
Domain/Application	Medical devices of all types (hardware, software/SaMD, IVD) & digital health; global regulatory scope.	General technology R&D; across industries (originated in aerospace; now multi-industry).	Cross-industry (software, IT, defense, aerospace, etc.) process improvement model.	Pharmaceuticals manufacturing (FDA/CDER pilot program).	Medical device manufacturing (voluntary improvement program by FDA CDRH and MDIC).	Medical devices worldwide (internationally harmonized QMS standard).
Maturity Levels	9 levels (Stage 1 through 9, with defined criteria for each).	9 levels (TRL 1-9).	5 levels (Maturity Levels 1-5).	No formal levels (qualitative assessment only).	5 maturity levels (uses CMMI maturity ratings, tailored to devices).	No maturity levels (pass/fail compliance standard).
Assessment Method	Self-assessment via structured checklist mapped to regulations, combined with algorithmic analysis (logistic model) to estimate approval probability.	Expert evaluation of development stage against defined TRL criteria (e.g. concept, lab prototype, clinical prototype, etc.).	Formal appraisal by certified CMMI assessors (evidence-based evaluation of processes against CMMI best practices).	FDA-led surveys and site visits assessing quality practices and culture (internally scored; pilot program).	Third-party CMMI-based appraisal for devices (team conducts on-site interviews and evidence review under FDA's Voluntary Improvement Program).	Third-party audit by a notified body or registrar to verify compliance with ISO 13485:2016 requirements.
Primary Output	Assigned RRL-QMS level (1-9) for the organization, plus an estimated probability of regulatory approval at that level (based on the model).	TRL score (1-9) indicating the technological readiness stage of the product.	CMMI maturity level certification (Level 1-5), or a detailed capability profile across process areas.	Qualitative maturity appraisal and feedback (an internal FDA score used to adjust oversight; results not public).	Appraisal report and a CMMI maturity level for the organization; used for internal improvement and potential regulatory incentives (e.g. reduced FDA inspections).	ISO 13485 compliance certificate (if audit is passed) or list of non-conformities (if not compliant).

Aspect	RRL-QMS (Regulatory Readiness Level - QMS)	TRL (Technology Readiness Levels)	CMMI (Capability Maturity Model Integration)	FDA-QMM (Quality Management Maturity)	MDDAP (MDIC Medical Device Discovery Appraisal Program)	ISO 13485 (QMS Standard)
Usage Context	Proposed for use by medical device companies to plan and track regulatory readiness; can inform investors and regulators for risk-based resource allocation (e.g. focusing support or oversight on lower-maturity organizations).	Used by R&D; teams and funders to gate product development progress and guide investment decisions; not specific to regulatory submissions.	Used by organizations to internally improve processes; also used by customers (e.g. government or prime contractors) to ensure supplier process quality (sometimes required in defense/IT contracts).	Used by FDA to encourage and guide manufacturers toward higher quality maturity; companies with higher maturity may receive reduced inspection frequency in pilot programs.	Used voluntarily by device firms to identify process impr ovements; part of FDA/MDIC's Case for Quality initiative. MDDAP results can lead to regulatory engagement benefits (e.g. streamlined inspections) for participants.	Mandatory QMS standard for market approval in many jurisdictions (compliance required for CE marking in EU, FDA approval in US, etc.); serves as the baseline for regulatory audits and inspections globally.

Table 3. Comparative characteristics of RRL-QMS vs. other maturity/readiness frameworks. RRL-QMS uniquely combines a staged regulatory compliance focus with a quantitative predictive engine, unlike technological maturity scales (TRL), general process models (CMMI), or recent quality maturity initiatives (QMM, MDDAP, ISO-based audits).

As seen in Table 3, RRL-QMS is distinct in its explicit regulatory focus and output. Whereas TRL and CMMI provide general measures of technology or process maturity, they are not tied to regulatory milestones and do not produce any estimate of approval probability. Programs like FDA-QMM and MDIC's MDDAP, while encouraging higher quality maturity, use qualitative appraisals and lack a standardized level-by-level progression or predictive analytics. In contrast, RRL-QMS introduces a nine-step roadmap mapped one-to-one to actual regulatory requirements (e.g. ISO 13485 clauses, MDSAP tasks) and couples it with a calibrated approval forecasting model. This integration of a quantitative prediction engine into a maturity framework is a novel contribution of RRL-QMS. Notably, RRL-QMS is able to indicate not just what maturity level an organization is at, but what that implies for their likelihood of regulatory success - a feature absent from all other frameworks.

Significance of Findings: The results highlight that an organization's QMS maturity (as measured by RRL-QMS) is a dominant determinant of regulatory outcomes. Our model suggests that improving one's RRL-QMS level has a larger impact on approval probability than even significant intrinsic factors like device risk class. This underlines the practical insight that investing in QMS development and regulatory preparedness can yield substantial returns in terms of success rates. In fact, the model attributes ~88% of outcome variance to RRL level alone, reinforcing the idea that regulatory readiness is often the decisive factor in approval outcomes when basic product viability is given. From a policy and management perspective, this finding supports a greater emphasis on early implementation of quality systems in the medtech start-up ecosystem. Teams with prior experience also fared much better, which is expected - experience likely captures tacit knowledge of regulatory processes - but interestingly, a high RRL can partly compensate for lack of prior experience by systematically guiding what needs to be

done.

Another important finding is the ability to quantify risk thresholds for decision-making. The decision curve analysis showed that our model can inform go/no-go decisions by quantifying the net benefit of delaying submissions for further improvements. For example, if an organization's current predicted chance of approval is only 50%, the model would support postponing the submission to implement corrective actions (since intervening in such "borderline" cases yields a positive net benefit). This kind of decision support is something static checklists or pass/fail audits do not provide. The introduction of the Stability Index further adds a layer of risk awareness: if a submission's success prediction is highly unstable (sensitive to minor data changes), that submission can be flagged as high-risk, prompting additional scrutiny or contingency plans. We found most cases had stability index ≥ 0.9, indicating the model's recommendations are generally robust; but for the minority of cases with lower stability, this metric serves as an early warning signal for borderline readiness.

Limitations and Future Work: The RRL-QMS framework, while comprehensive, represents a snapshot of current regulatory expectations and must be maintained as a "living" system that evolves over time. Regulations and best practices continue to develop (for example, new EU MDR interpretations or upcoming FDA QSR updates). We plan to institute a biennial review and versioning process: approximately every two years, the RRL-QMS criteria will be revisited and updated as needed, with new versions (e.g. RRL-QMS v2.0, v3.0) clearly delineated. This will ensure the framework stays aligned with the latest requirements and lessons learned. By committing to a scheduled update cycle, we aim to continuously incorporate industry feedback and improve the tool's predictive reliability.

Further large-scale, real-world validation of RRL-QMS is also planned. The current model's performance statistics (such as the 0.83 AUC and calibration) are based on simulations and a limited 10-case pilot. As the framework gains adoption in practice, ongoing data collection and outcome tracking will allow us to recalibrate and refine the model with real-world evidence. For example, we envision partnering with device accelerators or regulatory agencies to anonymously aggregate outcomes of submissions categorized by RRL-QMS level. This would enable continuously improving the model's accuracy and generalizability. It will be important to verify, in particular, that the odds-ratio per RRL level holds across different device categories and regions, and to adjust the model if certain levels prove too lenient or strict once empirical data accumulate. In summary, like any model intended for practical use, RRL-QMS will require periodic refinement and governance to maximize its value to the medical device ecosystem.

Encouraged by this study's results, we are also exploring extensions of the RRL framework to other domains. One immediate extension is to develop a variant for the pharmaceutical industry (tentatively, "RRL-P") aligned with guidelines such as ICH Q10 (Pharmaceutical Quality System). Another area is adapting the framework for emerging health technologies like AI/ML-based software as medical devices - here we would incorporate elements of Good Machine Learning Practice (GMLP) into the maturity criteria. The flexible, modular structure of RRL-QMS means new domains or specific requirements (e.g. data management for AI, or GMP for combination products) can be bolted on as additional checklist items or even new domains, without changing the overall ladder logic. Ultimately, we envision the RRL concept as a platform that could unify readiness assessment across various regulated product sectors, promoting a common "readiness language" analogous to how TRL unified technology readiness.

Finally, the content and findings of this work have been documented in multiple formats to balance openness with intellectual property protection. Table 4 maps the key concepts and results to the documents in which they are detailed or summarized: a full internal report (containing proprietary data and code), the present Zenodo preprint (public summary), and a separate abridged "Lite" publication. This matrix demonstrates that all major elements are traceably documented while sensitive implementation details remain protected.

Key Concept / Result	Internal Full Report (detailed)	Zenodo Preprint (this article)	“Lite” Abridged Article
Nine-level RRL-QMS framework (aligned to ISO 13485 & MDSAP)	Yes (comprehensive definitions)	Yes (overview provided)	Yes (summarized)
50-item domain-based checklist scoring mechanism	Yes (full checklist in appendix)	Yes (described, items withheld)	Yes (mentioned, not detailed)
Predictive logistic regression model (RRL + risk + experience)	Yes	Yes	Yes
Model performance metrics (≈1.7× odds per RRL, AUC ≈ 0.83, etc.)	Yes (full results)	Yes (reported in Results)	Yes (headline figures)
Decision curve analysis (net benefit of model-guided decisions)	Yes	Yes	Yes (briefly noted)
Global Sobol sensitivity analysis (variance ~88% by RRL)	Yes	Yes	Yes (implicit in text)
“Stability Index” for prediction robustness	Yes (defined, analyzed)	Yes (discussed)	No (not explicitly mentioned)
Real-world pilot study validation (10-case study)	Yes (included)	Yes (summarized)	No (not included)

Table 4. Documentation of key RRL-QMS concepts and results across source documents. Each checkmark indicates the concept/result is presented in the given document (the full internal technical report, the public preprint, and the “Lite” abridged article). Proprietary details (e.g. the full checklist and code) are confined to the internal report, while the public documents share the overarching framework and validation highlights.

Key Differentiators of the RRL-QMS Framework vs. Conventional “Gap-Checklists”

Conventional ISO 13485 or MDSAP “gap-analysis” templates are binary tools (“yes / no”) that indicate whether a document or procedure exists; they neither quantify regulatory success nor guide strategic decision-making. By contrast, RRL-QMS delivers three decisive advantages that elevate it from a checklist to a predictive decision framework:

Nine-Level Maturity Geometry A rigorously defined ladder (RRL-1 → RRL-9) captures the continuum from concept to post-market excellence, mapping each level to regulatory milestones (pre-submission, market access, surveillance). This phased architecture is absent from free templates, which lack a theory of staged progression.

Quantitative Predictive Engine A calibrated logistic-regression model translates checklist scores + contextual variables (device risk class, prior experience) into a probability of first-pass approval. The model is validated by ROC-AUC 0.83, Brier 0.166, global Sobol sensitivity, Monte-Carlo stability, and decision-curve analysis-features never reported for static checklists.

Risk-based Decision Support Output includes actionable thresholds (e.g., defer submission if predicted failure risk > 40 %); thus the framework informs go/no-go, budget allocation, and audit-readiness planning. Traditional checklists merely highlight gaps without quantifying their impact.

Dimension	Free Online Checklists	RRL-QMS Framework
Assessment Mode	Binary “Yes/No”	Binary inputs plus ordinal 9-level maturity

Dimension	Free Online Checklists	RRL-QMS Framework
Weighted Scoring	Rare / ad hoc	Yes - Delphi-derived weights
Probability of Approval	✗	✓ Logistic model & calibration
Sensitivity / Decision Analysis	✗	✓ Sobol, decision-curve, Monte-Carlo stability
Executable Codebase	✗	✓ Python package & reproducibility
SaaS / Licensing Readiness	Limited	✓ Commercial-grade IP & licensing

Table 5. Comparative features of generic gap-analysis checklists versus the RRL-QMS predictive framework. Sources for checklist limitations and RRL-QMS capabilities are detailed in the Lite summary version of this work.

Practical Applications of the RRL-QMS Framework

The Regulatory Readiness Level model is not merely a conceptual taxonomy; it is engineered for day-to-day decision-making across the digital-health ecosystem. Five primary stakeholder groups-and the concrete benefits they can expect-are summarised below.

(i) Manufacturers (SaMD & digital-health start-ups) Road-mapping and capital efficiency. From project inception, development teams can align sprint backlogs with the artefacts required at each RRL step, avoiding “catch-up” compliance late in the life-cycle. Early-stage firms, for example, can postpone costly clinical testing until they have demonstrably reached RRL-2 but must budget for QMS scaffolding in advance. Investors likewise obtain an intuitive progress indicator (“currently RRL-3, targeting RRL-5 in six months”) that de-risks regulatory diligence. Internally, every feature change is checked against its maturity impact-e.g., adding an AI module may regress the project to mid-RRL-2, triggering re-testing and documentation updates-allowing product and quality managers to co-estimate schedule and budget.

(ii) Certification & auditing bodies (e.g., EU Notified Bodies) Stage-gated audits and workload smoothing. Under the MDR regime, Notified Bodies (NBs) face resource bottlenecks because non-conformities surface late in the dossier review. A community-wide adoption of RRL would let NBs offer voluntary “pre-audit” checkpoints-say, at RRL-2 and again at RRL-3-so major gaps are closed before the final conformity assessment. Yearly surveillance audits could also benchmark whether a post-market product has slipped below its certified RRL, turning the ladder into an early-warning signal for quality erosion.

(iii) Regulators (FDA, MHRA, etc.) Resource prioritisation & policy analytics. In Q-Sub or Breakthrough interactions, FDA reviewers could ask sponsors to self-declare an RRL level, then corroborate it with objective evidence-allocating counsel where impact is greatest (e.g., RRL-3 submissions poised to finalise clinical protocols). Agencies can further aggregate national RRL statistics to track ecosystem maturity, much as they monitor inspection outcomes today.

(iv) Healthcare systems & end-users Transparency and procurement confidence. A simple label such as “RRL-5, FDA-cleared” conveys far more than a generic CE or 510(k) mark; it signals that the manufacturer maintains a mature post-market surveillance loop. Payers and hospital technology committees could incorporate minimum RRL thresholds into purchasing criteria, accelerating the adoption of demonstrably safer, more robust digital therapeutics.

(v) Future research & tooling Evidence generation and automation. Retrospective mapping of legacy SaMD products to RRL levels would test whether higher readiness indeed correlates with faster approvals or better post-market performance-an empirical validation path already under way. Online self-assessment portals and smart checklists can compute provisional RRL scores and recommend next actions, functioning as a “Regulatory Navigator” for innovators. Ultimately, RRL could merge with Market Readiness (MRL) indices to yield a 360-degree maturity metric covering regulatory, clinical-validation and commercial pillars.

In summary, the RRL-QMS framework converts regulatory-compliance complexity into an actionable, staged roadmap that underpins risk-based decision-making throughout the full medical-device life cycle, spanning stand-alone software, traditional hardware, in vitro diagnostics, and combination products.

Conclusion

In summary, we have introduced and validated the RRL-QMS framework, a multi-level maturity model tailored specifically for medical device regulatory compliance. The framework is structured into nine progressive levels aligned with international QMS standards (ISO 13485:2016) and the MDSAP audit process, providing a clear roadmap from initial regulatory planning to sustained post-market quality excellence. Using simulation-based modeling, pilot case studies, and advanced statistical analyses, we demonstrated that an organization's RRL-QMS level is an independent and powerful predictor of its regulatory approval success. Each incremental level in the framework nearly doubles the odds of approval, and RRL-QMS by itself explains an estimated ~88% of the variance in outcomes even after accounting for product risk and team experience. This indicates that regulatory readiness - as quantified by our framework - is often the decisive factor in approval outcomes.

By fusing a well-defined conceptual model of QMS maturity with a quantitatively validated predictive tool, RRL-QMS offers dual practical value for the medical device ecosystem. First, it serves as a strategic roadmap for product developers, providing a staged guide to navigate regulatory expectations and implement QMS milestones step by step. Teams can use it to identify gaps early, prioritize quality improvements, and objectively track their progress toward readiness for audits and submissions. Second, it functions as a decision-support tool for stakeholders such as investors, auditors, and regulators, who can use the RRL-QMS level and approval probability output to assess an organization's readiness and allocate resources accordingly. For example, venture investors might gauge the regulatory maturity of start-ups as part of due diligence, and regulators could triage submissions based on readiness scores - focusing attention and support where it's needed most.

Overall, the RRL-QMS framework establishes a foundation for standardized benchmarking of regulatory maturity. It enables more risk-informed, data-driven decisions throughout the device development lifecycle - from R&D; investment and pre-market planning to audit preparation and post-market surveillance. By making regulatory preparedness measurable, the framework and model can foster greater predictability and efficiency in bringing safe, effective medical technologies to global markets.

Importantly, this publication has shared the high-level framework and validation findings while safeguarding the intellectual property of the author. Full implementation details - including the complete 50-item RRL-QMS checklist, scoring algorithms, and validation code - are available upon reasonable request and only under a formal license agreement. This work is released under a CC BY-NC-ND 4.0 license, which permits non-commercial use with attribution but prohibits unauthorized derivatives. Users of the RRL-QMS model are thus encouraged to cite this original work and to engage via licensing for any detailed use or further development. By doing so, the community can build on a solid, validated foundation for regulatory readiness, while proper credit and control are maintained for this novel framework.

References

1. European Parliament and Council. Regulation (EU) 2017/745 of 5 April 2017 on medical devices (MDR). Official Journal of the European Union. 2017;L117.

2. European Parliament and Council. Regulation (EU) 2017/746 of 5 April 2017 on in vitro diagnostic medical devices (IVDR). Official Journal of the European Union. 2017;L117.

3. U.S. Food and Drug Administration (FDA); Center for Drug Evaluation and Research. Quality Management Maturity (QMM) Program. Silver Spring, MD: FDA; 2024.

4. Medical Device Innovation Consortium (MDIC). Assessing Regulatory Maturity: The Medical Device Discovery Appraisal Program (MDDAP). MDIC; 2022.

5. International Organization for Standardization (ISO). ISO 13485:2016 Medical devices-Quality management systems-Requirements for regulatory purposes. Geneva: ISO; 2016.

6. International Medical Device Regulators Forum (IMDRF). Medical Device Single Audit Program (MDSAP) - Annex 2: Audit model. IMDRF/MDSAP WG/N3; 2013.

7. Health Canada. Transition to the Medical Device Single Audit Program (MDSAP). Ottawa: Government of Canada; 2018.

8. Therapeutic Goods Administration (TGA). Use of MDSAP audit reports for TGA conformity assessment. Canberra: Australian Government Department of Health; 2019.

9. Saltelli A, Ratto M, Andres T, et al. Global Sensitivity Analysis: The Primer. Chichester: John Wiley & Sons; 2008.

10. Mankins JC. Technology readiness assessments: A retrospective. Acta Astronautica. 2009;65(9-10):1216-1223. doi:10.1016/j.actaastro.2009.01.026.

11. International Council for Harmonisation. ICH Q10: Pharmaceutical Quality System. Geneva: ICH; 2008.

12. Pharmaceuticals and Medical Devices Agency (PMDA). Utilization of MDSAP Audit Reports in Japan. Tokyo: PMDA; 2023.

13. National Medical Products Administration (NMPA). Good Manufacturing Practice for Medical Devices (Order No. 739). Beijing: NMPA; 2022.

14. World Health Organization. Global Model Regulatory Framework for Medical Devices including in vitro diagnostic medical devices. 2nd ed. Geneva: WHO; 2022.

15. Ries E. The Lean Startup. 1st ed. New York, NY: Crown Business; 2011.

16. Kuncheva LI. A stability index for feature selection. In: Proceedings of the International Conference on Artificial Neural Networks (ICANN); 2007:390-395. Prague, Czech Republic.

17. International Organization for Standardization (ISO). ISO/TR 24971:2020 Medical devices-Guidance on the application of ISO 14971. Geneva: ISO; 2020.

18. CMMI Institute. Capability Maturity Model Integration (CMMI) for Development, Version 2.0. Pittsburgh, PA: CMMI Institute; 2018.

19. U.S. Food and Drug Administration (FDA); Center for Devices and Radiological Health. Case for Quality - Voluntary Improvement Program Guidance. Silver Spring, MD: FDA; 2023.

20. Brier GW. Verification of forecasts expressed in terms of probability. Mon Weather Rev. 1950;78(1):1-3.