Risk-Based Quality Management

Practical Guide For Clinical Research Professionals

A. Andrianov PhD J. Proeve PhD

Second Edition

Copyright © 2026 A. Andrianov J. Proeve

All rights reserved.

ISBN: 9798243686594

DEDICATION

To our esteemed colleagues in the field of managing clinical trials, who dedicate immense time and effort to ensure patient safety, patient rights, and high-quality clinical trial data. Your commitment to excellence and continuous improvement has been a constant source of inspiration for enhancing your work while elevating the quality of your trials. This book is dedicated to you.

Artem & Johann

And to my beloved wife, Evgeniya. Your unwavering support and boundless inspiration have been the driving force behind every new achievement. This series of books is a testament to your love and encouragement.

Artem

CONTENTS

(….)

CASE STUDY – "THE QTL WE SWITCHED OFF"

Inspired by a true story shared by Marcin Makowski, Head of Centralized Monitoring at GSK; details anonymized and adapted for learning purposes.

1.1. Context: an ambitious inflammation study

Consider a large, global Phase III unblinded trial in a chronic inflammatory disease, such as rheumatoid arthritis. The study had two key ambitions: to demonstrate superior symptom control versus standard of care and to reduce dependence on rescue medication such as steroids and NSAIDs.

The design looked robust. The primary objective was improvement in an inflammatory score at week 24; a key secondary objective was reduction in the cumulative use of rescue medication. About 1,200 patients were planned across roughly 120 sites in several regions. The visit schedule was deliberately dense during the first 12 weeks to stabilize patients and then tapered thereafter.

From a quality-by-design perspective, the team checked the right boxes. They identified a Critical-to-Quality (CtQ) factor around adequate symptom control with minimal rescue-medication use and recognized that excessive rescue medication could mask the treatment effect, raise safety concerns linked to steroid toxicity, and complicate both regulatory and HTA interpretation. Rescue medication use was therefore selected as a study-level QTL. On paper, the logic was exemplary: CtQ → defined risk → QTL.

1.2. Designing the rescue-medication QTL

During protocol discussions and RACT workshops, the team translated this risk into a concrete parameter and threshold. The QTL parameter was defined as the mean number of rescue-medication days per patient, per treatment arm, over weeks 0–12. The threshold (for illustration) stipulated

that if the mean number of rescue days per patient in either arm exceeded around 10 days in that period, a cross-functional review would be triggered. That review was intended to bring together clinical operations, medical, statistical, and quality teams to analyze root causes—such as dosing, eligibility, adherence, or site practice—and to consider mitigations like clarifications, training, or small protocol adjustments within existing bounds.

The QTL was described in the SAP, referenced in the monitoring and RBQM plans, and recorded in the risk register, with rescue-medication use rated as "high" impact for both primary and key secondary endpoints. Structurally, this looked like a well-designed QbD/RBQM construct: clearly linked to CtQ, defined at trial level, and anchored in governance.

1.3. What actually happened: a QTL that was always red

As recruitment progressed, the central monitoring team began its regular data reviews. Very quickly, it became apparent that the mean number of rescue medication days per patient exceeded the predefined threshold. The QTL was breached at early data cuts, and then again, and again. For most of the first 12 weeks, the rescue medication QTL remained stubbornly red.

Three problems followed.

First came alert fatigue. Each governance meeting opened with the same message: "Rescue-medication QTL: red (again)." The signal felt constant and non-specific, discussions became repetitive, and the team struggled to convert this high-level alarm into targeted, practical actions. A QTL that never returns to green quickly stops acting as a trigger and becomes background noise.

Second, the QTL acquired the label of being "bad". Different stakeholders interpreted the persistent breach through their own lenses. Clinicians pointed out that patients were sicker than anticipated and argued that this level of rescue medication was quite typical for the population. Investigators, via CRA feedback, highlighted regional patterns

where rescue medication was used more liberally for reassurance or cultural reasons. Biostatisticians observed that while rescue-med use was high, it appeared to be balanced between arms, which made it less concerning from a comparative-efficacy standpoint. Gradually, the conversation shifted from "this QTL is telling us something important" to "this QTL is simply mis-specified."

Third came a quiet but decisive step: the team essentially switched off the QTL. In practice, this meant that no further QTL-breach alerts were raised for rescue-medication use, there was no formal recalibration or redefinition of the metric, and documentation was limited to meeting notes stating that the QTL was "not useful." The underlying metric continued to exist and would still be considered in the planned analyses; however, as a QTL and governance tool, it was quietly abandoned.

1.4. Downstream consequences: losing a story they could have owned

The trial ultimately achieved its primary objective, as evidenced by a statistically significant improvement in the inflammatory score. Rescue medication use, however, was higher than anticipated, and the differences between arms were less pronounced than initially hoped.

When the team prepared the submission and inspection-readiness materials, their earlier decision returned to haunt them. Regulators naturally inquired about what had happened when the rescue-medication QTL was breached, what actions had been taken in response, and why this QTL seemed to disappear from active oversight midway through the study.

Because the QTL had effectively been turned off without a formal redefinition or re-baselining, the resulting narrative looked reactive rather than proactive and left the organization exposed to questions about the maturity and discipline of its RBQM implementation. The problem was less the metric itself than the absence of a coherent story that showed continuous control over this CtQ factor.

Internally, the organization also lost an important learning opportunity. There was no structured review of whether the threshold had been unrealistic given the actual disease burden, whether the RACT had underestimated baseline rescue medication patterns in this indication and region mix, whether targeted site training or treatment algorithms could have influenced behavior, or whether specific pockets of extreme rescue use might have been addressed. None of these questions was captured in a formal "lessons learned" process; the QTL experience became an annoyance rather than a driver of better future design.

Finally, the team missed possible mitigations. Persistent high rescue-medication use early in the trial could have prompted immediate clarifications of rescue rules at investigator meetings, additional training on when to initiate, taper, or avoid rescue therapy, closer review of specific regions, investigators, or severity strata, or scenario modeling to understand the impact on the key secondary endpoint and on HTA arguments around steroid sparing. Instead, the trajectory effectively ran from "important risk to control" to "annoying red light" to "silenced," even though the underlying CtQ remained fully valid.

1.5. What a mature RBQM response would look like

This story is not primarily about a "bad QTL"; it is about governance and mindset. A more mature RBQM approach could have followed three deliberate steps.

The first would be to re-examine the QTL definition. Early on, the team could have inquired whether the threshold truly aligned with realistic historical data for this population and pattern of rescue medication practice, and whether the parameter itself accurately captured the right shape of risk. Perhaps total days of rescue medication were less informative than the number of high-dose steroid bursts, the proportion of patients with more than a certain number of rescue courses, or persistent high use after an initial adaptation period. If a QTL is always red, it may be poorly calibrated, but it may also be revealing a structural design issue—such as rescue rules that are too loose for the stated

objective—that needs to be addressed transparently. Treating the QTL as a hypothesis to refine, rather than a binary right-or-wrong construct, creates room for better decisions.

The second step would be to distinguish the signal from the pattern. Instead of a single global red/green indicator, the team could have examined trends over time, regional or country-level distributions, which would have required companion KRIs for the QTL, which are only available at the study-level, baseline-severity strata, and site or investigator clusters. It is often the case that the global mean exceeds the threshold, while most sites remain close to the target; some are moderate outliers, and a small group are extreme outliers. In that scenario, the QTL becomes an entry point to focused action, not an undifferentiated alarm.

The third step would be to re-baseline formally and document the change. If, after analysis, the team concluded that the original threshold was unrealistic, a mature RBQM response would involve prospectively updating the QTL definition or threshold, clearly recording the rationale (for example, underestimation of background rescue-med use or new external data), identifying the decision-makers across clinical, statistics, operations and quality, and explaining the impact on interpretation in the CSR and submission. The risk register, monitoring plan, and SAP would be updated to reflect the new definition. In short, QTLs are allowed to evolve, but they should never disappear silently.

1.6. Lessons learned: from "switch it off" to "learn from it"

The case illustrates a common RBQM pattern: a QTL that reveals more complexity than expected is quickly labeled "wrong" and turned off, instead of being refined and used as a learning tool.

Several lessons follow. QTLs should be seen as hypotheses based on the best knowledge available at the design stage, not as eternal truths. As real data accumulate, there is an expectation—not a failure—that they will be revisited and, if necessary, adjusted.

Governance is more important than perfection. An imperfect QTL that is regularly reviewed, recalibrated, and documented is more valuable than a seemingly perfect QTL that no one uses. When everything is red, it is not the users who are broken; it is the design. Persistent "red" states should trigger simplification of the metric, segmentation (for example, by region or severity), or recalibration, not silent deactivation.

The case also underscores that rescue medication is both a safety signal and an interpretability signal. High rescue use can undermine confidence in efficacy claims, even when arms appear balanced, because it may suggest inadequate disease control under real-world-like conditions and complicate HTA narratives about steroid sparing and long-term toxicity.

Finally, regulators expect a coherent story of control. When QTLs are defined in documents but abandoned in practice, it is reasonable for inspectors to question how seriously RBQM is taken and whether other stated controls are also more aspirational than operational. The underlying mindset shift is to stop fearing "difficult" QTLs; they often highlight precisely the areas where the organization has most to learn.

1.7. Practical considerations: designing and managing QTLs for rescue medication

This experience can be distilled into practical checkpoints that can be reused when designing rescue-medication QTLs or similar metrics.

Before the study, teams should pressure-test the realism of the planned thresholds against internal or published historical data and validate assumptions with investigators from high-use regions. The chosen parameter should be clinically interpretable—such as the frequency of high-dose bursts, the number of rescue courses, or failure to taper—and clearly connected to CtQs and key endpoints. Governance arrangements need to spell out how often the QTL will be reviewed, which forum owns the decisions, and how actions will be tracked and documented.

During the study, oversight should move beyond a single number and routinely examine time trends, geographic patterns, baseline-severity subgroups, and site- or investigator-level clustering. Responses should be proportional, with predefined action tiers: for example, a global trend above threshold could prompt an analytical deep dive, while regional or site outliers lead to targeted training, and evidence of harm or masked efficacy triggers review by the protocol steering committee. At fixed intervals, such as every three months, the team should explicitly reassess whether the QTL is oversensitive, insensitive, or still aligned with current understanding of the disease and trial conduct.

If the QTL appears "not to work" and there is pressure to switch it off, the first step should be to consider adjustments: re-baselining the threshold, redefining the parameter to better capture clinically meaningful patterns, or segmenting the metric to reduce noise. Any change should be prospectively documented with rationale and impact, reflected in the RACT, monitoring plan, and SAP, and agreed by multiple functions, not driven by a single team under operational stress.

1.8. Why this case matters

This case naturally links to broader discussions of QTLs, dynamic acceptable ranges, and RBQM governance. It shows that the hardest part of RBQM is not creating KRIs and QTLs, but living with them—especially when they highlight uncomfortable realities—and being disciplined in how they are refined.

When a QTL is always red, the right response is rarely to give up. More often, it is an invitation to deepen understanding of the disease, the protocol, and site behavior, and to turn what looks like a "bad QTL" into a better system for managing risk.

(….)

REFERENCES

Barnes, B., Stansbury, N., Brown, D., Garson, L., Gerard, G., Piccoli, N., Jendrasek, D., May, N., Castillo, V., Adelfio, A., Ramirez, N., McSweeney, A., Berlien, R., & Butler, P. (2021). Risk-Based Monitoring in Clinical Trials: Past, Present, and Future. *Therapeutic Innovation & Regulatory Science*, *55*. https://doi.org/10.1007/s43441-021-00295-8

Brown, M., PhD, A. H., Kilaru, R., Krohn, J., & PhD, Y. Z. (2023). *Harnessing the Power of Scientific Surveillance Applied Clinical Trials Online*. https://www.appliedclinicaltrialsonline.com/view/harnessing-the-power-of-scientific-surveillance

Buntz, B. (2025). Understanding protocol optimization in drug development. In *Drug Discovery and Development*. https://www.drugdiscoverytrends.com/complexitys-counterpoint-understanding-protocol-optimization/

Clinical Trial Eligibility Criteria for Patients with HIV: Guidance on Improving Inclusion in Non-HIV Trials. (2025). European AIDS Treatment Group. https://www.eatg.org/wp-content/uploads/2025/08/belong-draft-guidance.pdf

Clinical Trial Project Management: Proven Methods For PMOs. (2025). https://bestoutcome.com/knowledge-centre/clinical-trial-project-management/

Control Strategies and Good Clinical Practice (ICH E6 (R3))_20240708. (n.d.). *ACRO Transcelerate*.

Dirks, A., Florez, M., Torche, F., Young, S., Slizgi, B., & Getz, K. (2024). Comprehensive Assessment of Risk-Based Quality Management Adoption in Clinical Trials. *Ther Innov Regul Sci*, *58*(3), 520–527. https://doi.org/10.1007/s43441-024-00618-5

Dr. Baeyens, V. (2025). General Principles ICH E8 (R1) for Clinical Studies. In *Udemy*. Retrieved December 15, 2025, from https://www.udemy.com/course/general-considerations-ich-e8r1-for-clinical-studies/

Duggirala, H. J., Johnson, J. L., Tadesse, D. A., Hsu, C.-H., Norris, A. L., Faust, J., Walter-Grimm, L., & Colonius, T. (2025). Artificial intelligence and machine learning in veterinary medicine: A regulatory perspective on current initiatives and future prospects. *American Journal of Veterinary Research*, *86*(S1), S16–S21. https://doi.org/10.2460/ajvr.24.09.0285

Electronic Systems, Electronic Records, and Electronic Signatures in Clinical Investigations: Questions and Answers. (n.d.). *FDA*.

EudraLex - Volume 10 - Public Health - European Commission. (2025). https://health.ec.europa.eu/medicinal-products/eudralex/eudralex-volume-10_en

FDA. (2023). *A Risk-Based Approach to Monitoring of Clinical Investigations–Questions and Answers*.

FDA, C. for D. E. and R. (2025). *Conducting Clinical Trials With Decentralized Elements*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/conducting-clinical-trials-decentralized-elements

Fisher, A. (2023). ICH E6 (R3) Good Clinical Practice – MHRA Inspectorate. In *MHRA Inspectorate*. Retrieved January 2, 2026, from https://mhrainspectorate.blog.gov.uk/2023/05/26/ich-e6-r3-good-clinical-practice/

Grandinetti, C., Budwal-Jagait, M., Abid, H., Gebbia, E., Boley, E., Williams, L., Fisher, A., Marcus, L., Muldowney, L., Sellers, J., Wakelin-Smith, J., & Ayalew, K. (2025). Evolving Standards: Good Clinical Practice Insights from US FDA, MHRA UK, and Health Canada. *Clin Pharmacol Ther*, 10.1002/cpt.70120. https://doi.org/10.1002/cpt.70120

Gribben, J., Macintyre, E., Sonneveld, P., Doorduijn, J., Gisselbrecht, C., Jäger, U., Le Gouill, S., Rule, S., & Dreyling, M. (2020). Reducing Bureaucracy in Clinical Research: A Call for Action. *Hemasphere*, *4*(2), e352. https://doi.org/10.1097/HS9.0000000000000352

Guidance for Risk Management of Artificial Intelligence Systems. (2025). European Data Protection Supervisor. https://www.edps.europa.eu/system/files/2025-11/2025-11-11_ai_risks_management_guidance_en.pdf

Guideline on computerised systems and electronic data in clinical trials. (2023). EMA. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-computerised-systems-and-electronic-data-clinical-trials_en.pdf

High-level summary of the AI Act EU Artificial Intelligence Act. (n.d.). Retrieved January 4, 2026, from https://artificialintelligenceact.eu/high-level-summary/

Hurley, C., Shiely, F., Power, J., Clarke, M., Eustace, J. A., Flanagan, E., & Kearney, P. M. (2016). Risk based monitoring (RBM) tools for clinical trials: A systematic review. *Contemporary Clinical Trials*, *51*, 15–27. https://doi.org/10.1016/j.cct.2016.09.003

ICH E6 Asset Library. (2025). In *TransCelerate*. https://www.transceleratebiopharmainc.com/assets/ich-e6-asset-library/

ICH E6(R3) Explained: Key Changes to GCP Guidelines. (n.d.). In *IntuitionLabs*. Retrieved December 15, 2025, from https://intuitionlabs.ai/articles/ich-e6-r3-gcp-guidelines-2026

Jesse Gelsinger. (2025). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Jesse_Gelsinger&oldid=1322554399

Journot, V., Pignon, J.-P., Gaultier, C., Daurat, V., Bouxin-Métro, A., Giraudeau, B., Preux, P.-M., Tréluyer, J.-M., Chevret, S., Plättner, V., Thalamas, C., Clisant, S., Ravaud, P., & Chêne, G. (2011). Validation of a risk-assessment scale and a risk-adapted monitoring plan for academic clinical research studies — The Pre-Optimon study. *Contemporary Clinical Trials*, *32*(1), 16–24. https://doi.org/10.1016/j.cct.2010.10.001

Klatte, K., Pauli-Magnus, C., Love, S., Sydes, M., Benkert, P., Bruni, N., Ewald, H., Arnaiz Jimenez, P., Bonde, M. M., & Briel, M. (2019). Monitoring strategies for clinical intervention studies. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.MR000051

Manfrin, A., Lee, K., Cacou, C., Pound, J., Harnden, A., Raine, J., & Pirmohamed, M. (2026). Evaluation of the Medicines and Healthcare products Regulatory Agency's introduction of a risk-proportionate approach for clinical trials: An analysis of 4617 applications assessed between September 2023 and August 2024. *British Journal of Clinical Pharmacology*, *n/a*(n/a). https://doi.org/10.1002/bcp.70308

Niangoran, S., Barbieri, A., Badje, A., Journot, V., Kouame, G., Marcy, O., Anglaret, X., & Alioum, A. (2025). A New Centralized Statistical Monitoring Method for Detecting Atypical Distribution of Qualitative Variables in Multicenter Randomized Controlled Trials. *Statistics in Biopharmaceutical Research*, *17*(3), 479–487. https://doi.org/10.1080/19466315.2024.2404631

Reddy, S. (2025). Global Harmonization of Artificial Intelligence-Enabled Software as a Medical Device Regulation: Addressing Challenges and Unifying Standards. *Mayo Clinic Proceedings: Digital Health*, *3*(1), 100191. https://doi.org/10.1016/j.mcpdig.2024.100191

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). (2024). http://data.europa.eu/eli/reg/2024/1689/oj

Research, C. for D. E. and. (2022). *E8(R1) General Considerations for Clinical Studies*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e8r1-general-considerations-clinical-studies

Riddle, J. (2024). Understanding the Impact of the New EU Artificial Intelligence Act on Clinical Research. In *Advarra*. https://www.advarra.com/blog/understanding-the-impact-of-the-new-eu-artificial-intelligence-act-on-clinical-research/

Risk-Adapted Approach to Clinical Trials and Risk Assessments. (n.d.). In *GOV.UK*. Retrieved December 17, 2025, from https://www.gov.uk/government/publications/risk-adapted-approach-to-clinical-trials-and-risk-assessments

Sibbald, B. (2001). Death but one unintended consequence of gene-therapy trial. *CMAJ*, *164*(11), 1612. https://pmc.ncbi.nlm.nih.gov/articles/PMC81135/

Stansbury, N., Barnes, B., Adams, A., Berlien, R., Branco, D., Brown, D., Butler, P., Garson, L., Jendrasek, D., Manasco, G., Ramirez, N., Sanjuan, N., Worman, G., & Adelfio, A. (2022). Risk-Based Monitoring in Clinical Trials: Increased Adoption Throughout 2020. *Therapeutic Innovation & Regulatory Science*, *56*, 1–8. https://doi.org/10.1007/s43441-022-00387-z

Stansbury, N., Branco, D., McDavid, C., Stewart, J., Surdam, K., Olson, N., Perry, J., Liska, J., Phillips, L., Coogan, A., Adelfio, A., & Garson, L. (2025). Risk-Based Quality Management: A Case for Centralized Monitoring. *Ther Innov Regul Sci*, *59*(2), 199–210. https://doi.org/10.1007/s43441-024-00719-1

Teodoro, D., Naderi, N., Yazdani, A., Zhang, B., & Bornet, A. (2025). A scoping review of artificial intelligence applications in clinical trial risk assessment. *Npj Digit. Med.*, *8*(1), 486. https://doi.org/10.1038/s41746-025-01886-7

The TransCelerate Model Approach. (n.d.). In *TransCelerate*. Retrieved December 15, 2025, from https://www.transceleratebiopharmainc.com/rbminteractiveguide/how-does-clinical-trial-site-monitoring-work-under-a-risk-based-monitoring-approach/the-transcelerate-model/

TransCelerate - Risk Based Monitoring - Clinical Trials. (n.d.). In *TransCelerate*. Retrieved December 15, 2025, from https://www.transceleratebiopharmainc.com/initiatives/risk-based-monitoring/

TUFTS. (n.d.). *Impact Reports : Tufts CSDD : Tufts University*. Tufts Center for the Study of Drug Development. Retrieved January 2, 2026, from https://csdd.tufts.edu/publications/impact-reports

Updated document - ICH Q9 Quality Risk Management - Public Health. (n.d.). Retrieved December 17, 2025, from https://health.ec.europa.eu/latest-updates/updated-document-ich-q9-quality-risk-management-2023-10-13_en

US FDA qualifies first AI tool to help speed liver disease drug development. (2025). *Reuters*. https://www.reuters.com/business/healthcare-pharmaceuticals/us-fda-qualifies-first-ai-tool-help-speed-liver-disease-drug-development-2025-12-09/

US FDA recommends steps to improve diversity in clinical trials. (2024). *Reuters*. https://www.reuters.com/business/healthcare-pharmaceuticals/us-fda-issues-guidance-drugmakers-diversity-studies-2024-06-26/

Viron, S., Trotta, L., Steijn, W., Young, S., & Buyse, M. (2024). Does Central Statistical Monitoring Improve Data Quality? An Analysis of 1,111 Sites in 159 Clinical Trials. *Therapeutic Innovation & Regulatory Science*, *58*. https://doi.org/10.1007/s43441-024-00613-w

Whitehead, M., Suprin, M., Mistree, T., Kearns, M. M., Marini, G., Goffe, C., Pillwein, M., & Abdul-Shukkoor, V. (2024). The Renovation of Good Clinical Practice: A Framework for Key Components of ICH E8. *Ther Innov Regul Sci*, *58*(2), 303–310. https://doi.org/10.1007/s43441-023-00561-x

Willigers, B. J. A., Wiesiolek, A. E., Potter, B., Pszczółkowska, A., Kristensson, C., Presz, V., Saraiva, G. L., & Psallidas, I. (2025). Development of a protocol complexity tool: A framework designed to stimulate discussion and simplify study design. *BMC Med Res Methodol*, *25*, 205. https://doi.org/10.1186/s12874-025-02652-9