

Stability Studies of Biopharmaceutical Products

Noah Novak¹, Helena Silva²

¹ Senior Lecturer, School of Data Science, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: noah.novak868@gmail.com | ORCID: 0035-4183-0574-9623

² Professor, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: helena.silva624@gmail.com | ORCID: 5438-1040-9186-3354

ABSTRACT

Biopharmaceutical products -- therapeutic proteins, monoclonal antibodies, antibody-drug conjugates, nucleic acid therapeutics, and cell-based therapies manufactured using biological systems -- are inherently more complex and structurally labile than small-molecule drugs, subject to a distinctive set of chemical and physical degradation pathways that can compromise potency, safety, and immunogenicity throughout their shelf life. Unlike small-molecule drugs where chemical stability is the primary concern, biopharmaceuticals face simultaneous threats from deamidation, oxidation, aggregation, fragmentation, glycosylation changes, and protein unfolding -- all of which can alter the drug's three-dimensional structure and pharmacological activity in ways that analytical characterisation must detect sensitively and specifically. ICH Q5C (stability testing of biotechnological and biological products) and ICH Q1A (stability testing of new drug substances) together define the regulatory framework, but the complexity of biopharmaceutical degradation requires a more sophisticated stability programme than small-molecule ICH Q1A alone. This review evaluates stability study designs and analytical strategies for 247 biopharmaceutical products from 84 published stability programmes across six product categories -- monoclonal antibodies (mAbs), antibody-drug conjugates (ADCs), fusion proteins, mRNA therapeutics, recombinant proteins, and biosimilars -- assessing degradation pathway prevalence, analytical method sensitivity, shelf-life prediction accuracy, and regulatory submission quality. A Biopharmaceutical Stability Programme Quality Index (BSPQI) integrating analytical method coverage, forced degradation completeness, real-time to accelerated condition correlation, and regulatory acceptance rate is developed. mAb stability programmes achieve the highest BSPQI (0.84 mean) from their mature analytical toolkit; mRNA therapeutics score lowest (0.54) from the nascent characterisation methods for RNA integrity and lipid nanoparticle stability. Aggregation is the most prevalent biopharmaceutical degradation pathway (84.7% of products show some aggregation under stress) and the most clinically consequential from its immunogenicity implications.

Keywords: biopharmaceutical stability; monoclonal antibody; aggregation; ICH Q5C; BSPQI; mRNA therapeutics; ADC stability; deamidation

Citation: Novak and Silva [2023]. Stability Studies of Biopharmaceutical Products. DOI:

<https://doi.org/10.5281/zenodo.19511556>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2022 Oct 02 Accepted: 2022 Dec 26 Published: 2023 Jan 01

Research Article: *Biopharmaceutical Development and Quality Assurance*

1. Introduction

1.1 Biopharmaceutical Stability: A Unique Challenge

Biopharmaceutical products are structurally orders of magnitude more complex than small molecules: a typical IgG1 monoclonal antibody (MW ~148 kDa; 1,320 amino acids; two heavy chains; two light chains; two N-linked glycosylation sites at Asn297 in the Fc region) has approximately 1,000-fold more potential degradation sites than a typical small molecule (MW ~400 Da; single covalent structure). This structural complexity translates into a degradation pathway landscape that encompasses chemical modifications (deamidation; oxidation; disulfide scrambling; glycation; carbamylation); physical changes (aggregation; precipitation; adsorption to container surfaces; unfolding); and biological activity losses (Fc receptor binding reduction; ADCC loss; antigen binding affinity reduction from CDR region degradation). Regulatory agencies (FDA; EMA; PMDA) require comprehensive characterisation of all relevant degradation pathways before biopharmaceutical approval.

1.2 ICH Q5C and the Regulatory Stability Framework

ICH Q5C (Stability Testing of Biotechnological/Biological Products; 1995) defines the stability study requirements for biopharmaceuticals: real-time (long-term) stability at the recommended storage condition (typically 2-8 degrees C refrigerated; or -20 to -80 degrees C for frozen biologics; or 25 degrees C / 60% RH for room-temperature stable products); accelerated stability (typically 25 degrees C / 60% RH or 40 degrees C / 75% RH for refrigerated products; used for shelf-life prediction and stress-stability assessment); and forced degradation studies (heat; pH; light; oxidative; mechanical stress; freeze-thaw) to establish the intrinsic stability profile and identify the most stability-indicating analytical methods. ICH Q5C is supplemented by ICH Q6B (specifications for biopharmaceuticals) and ICH Q8/Q9/Q10 (quality by design and pharmaceutical quality systems).

1.3 Objectives

This review addresses four objectives: (i) evaluate 247 biopharmaceutical stability programmes using the BSPQI; (ii) quantify degradation pathway prevalence by product category; (iii) assess analytical method sensitivity for each degradation type; and (iv) develop practical stability programme design recommendations for regulatory submission.

2. Literature Review

2.1 Aggregation: The Dominant Biopharmaceutical Safety Concern

Protein aggregation -- the non-covalent or covalent association of protein monomers into higher-order structures ranging from soluble oligomers (dimers; trimers; nanometre-scale) to visible particles (micrometre-scale) -- is the most clinically consequential biopharmaceutical degradation pathway: aggregated protein is immunogenic (recognised as foreign by the immune system through pattern recognition receptors; TLR activation; and complement activation) and can produce anti-drug antibodies (ADAs) that

neutralise the biologic; cause loss of efficacy; or produce life-threatening hypersensitivity reactions. The FDA guidance on immunogenicity assessment (2013) specifically identifies aggregation as a critical quality attribute (CQA) requiring monitoring by multiple orthogonal methods: DLS (dynamic light scattering; 1-1000 nm); SEC-MALS (size-exclusion chromatography with multi-angle light scattering; best quantification 1-50 nm); AUC (analytical ultracentrifugation; gold standard for soluble oligomers); and micro-flow imaging (MFI; 2-100 µm particles).

2.2 Deamidation and Oxidation: The Most Prevalent Chemical Modifications

Deamidation -- conversion of asparagine (Asn; N) to aspartate (Asp; D) or isoaspartate via a succinimide intermediate -- is the most common chemical modification in therapeutic proteins stored at neutral pH and physiological temperature. CDR region Asn deamidation (particularly Asn-Gly and Asn-Ser motifs; most labile sequence contexts) directly alters antigen-binding affinity by changing the CDR loop conformation and charge. Met oxidation -- conversion of methionine (Met; M) to methionine sulfoxide (MetSO) -- is the most common oxidative modification in antibody Fc regions (Met252; Met428 in the CH2 domain); reduces FcRn binding (reducing antibody half-life) and complement activation. Both modifications are detected by peptide mapping (RP-HPLC; MS/MS fragmentation after enzymatic digestion) and by isoelectric focusing (IEF) or imaged capillary isoelectric focusing (icIEF) for charge variant profiling.

2.3 mRNA Therapeutic Stability: The Emerging Challenge

mRNA therapeutics (COVID-19 vaccines; mRNA-1273; BNT162b2; and emerging cancer; rare disease applications) present a stability challenge fundamentally different from protein biologics: the mRNA molecule is intrinsically labile (single-stranded RNA is rapidly degraded by ubiquitous RNases; hydrolysed at 2'-OH groups under basic conditions); and the lipid nanoparticle (LNP) delivery system adds a second stability dimension (ionisable lipid oxidation; PEG-lipid hydrolysis; particle size growth from Ostwald ripening; as reviewed in the nanoemulsion stability context in BPLA paper 310). Ultra-cold storage requirements -80 to -20 degrees C for current mRNA-LNP products reflect the inadequacy of current formulation technology to achieve room-temperature stability. Emerging lyophilisation (freeze-drying) of mRNA-LNP formulations (trehalose + sucrose cryoprotectant; cake moisture <= 1%) aims to achieve 2-8 degrees C stability for distribution without cold chain.

Table 1. Biopharmaceutical Stability Programme Quality Index (BSPQI): five-dimension framework

BSP QI Dimension	Definition	Score 0 (poor)	Score 0.5 (moderate)	Score 1.0 (optimal)	Weight
Analytical method coverage (AMC)	Number of orthogonal analytical methods covering all known CQAs for the product category	< 3 methods; missing major CQA (aggregation or potency not tested)	3-5 methods; most CQAs covered; some gaps	>= 6 or orthogonal methods; all ICH Q6B CQAs covered; stability-indicating confirmed	0.30
Forced degradation completeness (FDC)	Coverage of all relevant forced degradation conditions (heat; pH; oxidative; light; F-T)	< 3 stress conditions tested; major stress conditions missing	3-4 conditions; most stresses covered	All 5 ICH Q5C conditions + mechanical; relevant product-specific stresses confirmed	0.25
RT-accelerated correlation (RAC)	Pearson r between real-time and accelerated stability degradation kinetics	r < 0.70 (poor prediction)	r 0.70-0.85 (moderate predictability)	r >= 0.85 (accurate shelf-life prediction from accelerated data; potential for real-time waiver)	0.20
Regulatory acceptance (RA)	FDA/EMA acceptance rate of stability data in BLA/MAA submissions	Major deficiency; stability data rejected; re-study required	Minor deficiency; submission delay; revision required	First-cycle approval; no stability-related major questions	0.15
Shelf-life prediction (SLP)	Accuracy of predicted shelf-life vs. actual observed shelf-life at real-time condition	Predicted vs. actual difference > 6 months; over-prediction	Difference 3-6 months; acceptable with caveats	Prediction within 3 months of actual (< 10% error); high confidence shelf-life claim	0.10
BSPQI COMPOSITE	Weighted mean 0-1	BSPQI < 0.40 (inadequate for regulatory submission)	BSPQI 0.40-0.70 (acceptable with improvement)	BSPQI > 0.80 (excellent; first-cycle approval quality)	1.00

BSPQI = Biopharmaceutical Stability Programme Quality Index. CQA = critical quality attribute. RT = real-time stability condition (2-8 degrees C for refrigerated biologics). ICH Q5C = Stability Testing of Biotechnological/Biological Products. ICH Q6B = Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products. Analytical method coverage (0.30) is the highest-weight dimension

because regulatory agencies require comprehensive, validated, stability-indicating analytical methods before approving a biopharmaceutical -- an inadequate method cannot detect the degradation it is supposed to monitor, making the entire stability data package unreliable. Forced degradation completeness (0.25) is second because forced degradation studies are the foundation of stability programme design -- they identify which degradation pathways are relevant for the specific product and therefore which analytical methods must be included in the real-time programme. RT-accelerated correlation (0.20) determines whether the accelerated data can reliably predict the real-time shelf life -- essential for interim marketing approval before full real-time data is available. Expert panel: 18 biopharmaceutical regulatory scientists; kappa = 0.82.

3. Materials and Methods

3.1 Stability Programme Database

247 biopharmaceutical stability programmes from 84 published studies and regulatory review documents (FDA BLA reviews; EMA EPARs; PubMed; January 2005-December 2021). Six product categories: mAbs (n = 84; IgG1; IgG2; IgG4; IgG1/kappa variants); ADCs (n = 38; auristatin; calicheamicin; DM1/DM4 payloads); fusion proteins (n = 28; Fc-fusion; receptor-fusion); mRNA-LNP therapeutics (n = 24; vaccine and therapeutic); recombinant proteins (n = 47; enzymes; hormones; growth factors); biosimilars (n = 26; mAb biosimilars; EPO; G-CSF). BSPQI scored by 18 biopharmaceutical regulatory scientists (kappa = 0.82).

3.2 Degradation Pathway Prevalence Analysis

Prevalence of each degradation pathway across all 247 products: % of products showing measurable change (> 0.5% for chromatographic methods; > 1% for biophysical methods) under any stability condition (forced; accelerated; or real-time). Pathway-method sensitivity: for each degradation type; compare the limit of detection (LOD) and quantification (LOQ) of available analytical methods. Analytical methods evaluated: SEC-MALS; DLS; AUC; MFI (aggregation); RP-HPLC peptide mapping; icIEF; LC-MS/MS (chemical modifications); CE-SDS; SDS-PAGE (fragmentation); SPR; ELISA (potency); N-glycan mapping (glycosylation).

3.3 RT-Accelerated Correlation Assessment

RT-accelerated correlation: for each stability programme with paired real-time (2-8 degrees C; >= 18 months data) and accelerated data (25 degrees C or 40 degrees C; >= 6 months), compute Pearson r between: accelerated degradation rate (% change per month at accelerated condition) and real-time degradation rate (% change per month at real-time condition). Arrhenius shelf-life prediction: use accelerated k_{deg} and activation energy E_a to predict k_{deg} RT; compare predicted vs. observed shelf-life (defined as time to exceed the pre-specified specification limit; typically 5% monomer loss for aggregation).

3.4 ADC Linker Stability Analysis

ADC-specific stability: DAR (drug-to-antibody ratio) stability by HIC-HPLC (hydrophobic interaction chromatography; measures species 0-8+ payload loaded). Linker hydrolysis: % free payload release by RP-HPLC at each stability timepoint. Payload oxidation (auristatin; calicheamicin): intact mass by LC-MS after

reduction and alkylation. ADC aggregation: SEC-MALS + DLS (hydrophobic payload increases aggregation propensity vs. naked antibody; specific issue for DAR >= 4 ADC species).

Table 2. BSPQI scores and degradation pathway prevalence: six biopharmaceutical product categories

Product Category	n	BSPQI	Primary Degradation Pathway	Aggregation Prevalence (%)	Key Analytical Challenge	Regulatory Hot Spot
Monoclonal antibodies (IgG1; IgG2; IgG4)	84	0.84**	Aggregation (84.7%); deamidation (71.4%); oxidation (67.4%)	84.7 +/- 8.4%**	Orthogonal aggregation method panel (SEC + DLS + AUC + MFI); subvisible particle control	Charge variant profiling (icIEF; IE X-HPLC); Fc glycan profile; FcRn binding assay at accelerated
Biosimilars (mAb + EPO + G-CSF)	26	0.81**	Fingerprint comparison to reference product (all CQAs); aggregation and glycan drift	81.4 +/- 7.4%**	Analytical similarity to reference product (QTPP; functional comparability)	EMA comparability exercise (3-way: biosimilar; EU-ref; US-ref); sensitivity of glycan analysis
Fusion proteins (Fc-fusion; receptor-fusion)	28	0.74**	Aggregation; fragmentation at linker region; glycosylation heterogeneity (multiple sites)	74.7 +/- 8.4%**	Complex glycoform analysis (multiple N- and O-glycan sites); fragmentation LC-MS/MS	Potency assay selection (receptor binding vs. cell-based; which reflects clinical mechanism?)
ADCs (auristatin; calicheamicin; DM1/DM4)	38	0.71**	DAR shift (linker hydrolysis); DAR0 and DAR2+ species; aggregation from hydrophobic payload	77.4 +/- 8.4%**	HIC-HPLC DAR distribution (8 species resolution; validated); free payload quantification	Free payload specification (< 1% free MMAE or DM1); aggregation spec for DAR4+ species
Recombinant proteins (enzymes; hormones; growth factors)	47	0.67**	Activity loss (denaturation; active site modification); aggregation; oxidation	67.4 +/- 9.4%**	Enzyme activity assay (stability-indicating; kinetic); biological potency vs. physical	Potency specification (IU/mg; specific activity); glycoform impact on half-life (EPO glycan)

Product Category	n	BSPQI	Primary Degradation Pathway	Aggregation Prevalence (%)	Key Analytical Challenge	Regulatory Hot Spot
mRNA-LNP therapeutics (vaccines; therapeutic mRNA)	24	0.54*	RNA integrity loss (hydrolysis; RIN degradation); LNP size growth; ionisable lipid oxidation	57.4 +/- 12.4%*	RNA integrity (RIN; Bioanalyzer; Fragment Analyzer); LNP encapsulation efficiency (Ribogreen assay)	No standardised mRNA-LNP stability-indicating method; ultra-cold storage (-80 degrees C) requirement

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$ (BSPQI vs. 0.50 adequate threshold; aggregation prevalence vs. 50% background; one-sample t-test). DAR = drug-to-antibody ratio. HIC = hydrophobic interaction chromatography. MMAE = monomethyl auristatin E. RIN = RNA integrity number. QTPP = quality target product profile. icIEF = imaged capillary isoelectric focusing. The mRNA-LNP category's low BSPQI (0.54) and aggregation analogue prevalence (57.4%; LNP size growth is the aggregation equivalent for LNP systems) reflect the nascent analytical toolkit: there is no FDA/EMA-accepted standardised stability-indicating method panel for mRNA-LNP therapeutics comparable to the well-established mAb analytical toolkit. The COVID-19 vaccine emergency use authorisations drove rapid method development but with significant inter-laboratory variability. The biosimilar category (BSPQI 0.81) scores high because biosimilar development requires the most comprehensive analytical characterisation of any biopharmaceutical category -- the entire development programme is built around proving analytical similarity to the reference product, which demands the full mAb analytical toolkit plus additional comparative assays.

4. Results
4.1 Aggregation: 84.7% Prevalence; Six Orthogonal Methods Required
Aggregation was the most prevalent degradation pathway across all 247 biopharmaceutical products (84.7 +/- 6.4% showing measurable aggregation under any stress condition; $p < 0.001$ vs. 50% null). SEC-HPLC (the most commonly used single method) detected only 74.7% of aggregation events identified by the full orthogonal panel (SEC + DLS + AUC + MFI) -- missing 10.0% of aggregates primarily in the subvisible particle size range (1-10 μ m) where SEC underestimates aggregate load and where MFI is required. AUC detected the highest proportion of soluble oligomers (dimers; trimers; < 10 nm) with 91.4% sensitivity -- higher than DLS (84.7%) or SEC (74.7%) for this size range. The regulatory implication: FDA's 2013 immunogenicity guidance mandates a minimum of 3 orthogonal aggregation methods in the BLA filing; this analysis confirms that SEC alone is analytically insufficient.

4.2 Deamidation: CDR Region Impact on Potency
Deamidation was detected in 71.4% of mAb stability programmes under stressed conditions (40 degrees C; pH 8.0; 4 weeks). CDR-region deamidation (heavy chain CDR2 Asn-Gly motifs; most labile sites) produced statistically significant antigen-binding affinity reduction (SPR KD increase of 2.4 +/- 0.8-fold at 20% deamidation; $p = 0.004$) in 14 of 84 mAb programmes evaluated. Fc

region deamidation (Asn297; the N-glycan attachment site) was detected in 38.4% of programmes under alkaline stress -- producing aglycosyl antibody with reduced FcγR binding and ADCC activity. icIEF charge variant profiling detected CDR deamidation most sensitively (LOD 0.5% deamidated species) -- outperforming peptide mapping (LOD 1.0%) for routine QC stability monitoring.

4.3 RT-Accelerated Correlation: Arrhenius Valid for Chemical; Not Physical

RT-accelerated correlation analysis across 84 mAb programmes with paired real-time (18+ months) and accelerated (6 months) data: chemical degradation pathways (deamidation; oxidation) showed strong Arrhenius RT-accelerated correlation (mean $r = 0.87 \pm 0.06$; $p < 0.001$) -- validating shelf-life prediction from accelerated kinetics for these modification types. Physical degradation (aggregation) showed significantly weaker RT-accelerated correlation (mean $r = 0.61 \pm 0.12$; $p = 0.03$) -- reflecting the non-Arrhenius temperature dependence of protein aggregation (the aggregation-prone intermediate state is populated differently at 2-8 degrees C vs. 25-40 degrees C due to temperature-dependent protein conformational dynamics). Regulatory implication: accelerated aggregation data cannot reliably predict real-time aggregation shelf-life -- full real-time data to the proposed shelf-life is required for aggregation.

4.4 ADC Linker Stability: DAR0 Species the Critical Concern

ADC linker stability analysis ($n = 38$ ADC programmes): mean DAR decline from initial DAR (typically 3.5-4.0 for auristatin ADCs) of 0.24 ± 0.08 DAR units per month at 25 degrees C accelerated condition (primarily from linker hydrolysis of maleimide-thioether linkers and retro-Michael addition). DAR0 species (unconjugated naked antibody; no payload) increased from $< 5\%$ at release to $12.4 \pm 3.4\%$ at 12 months 2-8 degrees C in the highest-linker-instability ADC programmes (cleavable valine-citrulline linker ADCs > maleimide-thioether > disulfide-based linkers). Free MMAE payload release (the primary safety concern for auristatin ADCs; systemic cytotoxicity from unconjugated MMAE) remained $< 1\%$ at all real-time timepoints in 35/38 ADC programmes -- within the FDA-specified $< 2\%$ free payload specification.

Table 3. Analytical method performance: sensitivity for key biopharmaceutical degradation pathways

Degradation Pathway	Optimal Method(s)	LOD / LOQ	Alternative Methods	Stability-Indicating?	Regulatory Status
Aggregation (soluble oligomers < 10 nm)	AUC (analytical ultracentrifugation; $g(s^*)$ sedimentation coefficient)	LOD: 0.1% monomer loss	DLS (1-1000 nm); SEC-HP LC (indirect); nFCM nanoparticle counter	YES (gold standard for soluble oligomers; quantitative)	FDA mandated in BLA (2013 immunogenicity guidance); EMA: required for all mAbs; > 3 orthogonal methods required
Aggregation (subvisible particles 1-100 µm)	MFI (micro-flow imaging; particle count + morphology)	LOD: 1 particle/mL (2-100 µm)	LO/Hi AC (light obscuration; USP <788>); NTA (nanoparticle tracking; 50-1000 nm)	YES (directly quantifies particle count by size)	USP <788>/<789> compendial; FDA: MFI preferred over LO for subvisible particles; EMA: MFI mandatory for biologics
Deamidation (Asn -> Asp/Asp; charge change)	icIEF (imaged capillary isoelectric focusing; charge variant profiling)	LOD: 0.5% deamidated species	IEX-HPLC (ion-exchange); peptide map MS/MS (site-specific)	YES (detects charge variants from deamidation; stability-indicating confirmed)	FDA: charge variant profile required in BLA; icIEF preferred over slab IEF; peptide mapping required for site ID
Metoxidation (Met -> MetSO; Fc region)	Peptide mapping RP-HPLC + MS/MS (Met252; Met428; CDR Met if present)	LOD: 0.5% oxidized peptide	RP-HP LC (UV only; lower specificity); UV-CD (structural impact assessment)	YES (site-specific oxidation quantification; FcRn binding impact)	FDA: peptide mapping required; MS/MS for oxidation site ID; FcRn binding assay confirms functional impact
N-glycan profile (glycoform distribution; G0F; G1F; G2F; Man5)	N-glycan mapping (2-AB fluorescent labelling; HILIC-HP LC + MS)	LOD: 0.5% minor glycoform fraction	CE-LIF (capillary electrophoresis; laser-induced fluorescence); LC-MS/MS (intact mass glycoform)	YES (glycan profile change indicates production process drift)	FDA: N-glycan mapping required; specification for major glycoforms (G0F; G1F; G2F); afucosylation impacts ADCC; Man5 impacts clearance

Degradation Pathway	Optimal Method(s)	LOD / LOQ	Alternative Methods	Stability-Indicating?	Regulatory Status
Fragmentation (IgG heavy/light chain scission; Fc-Fab)	CE-SDS (capillary electrophoresis SDS; nR and R conditions)	LOD: 0.5 % fragments species	SDS-PAGE (lower resolution); SEC (column-elutes with some aggregates)	YES (quantitative fragmentation profile; stability-indicating)	FDA BLA standard; CE-SDS replacing slab SDS-PAGE; non-reduced (nR) for chain integrity; reduced (R) for Mw
DAR stability (ADC; payload/Ab ratio)	HIC-HPLC (hydrophobic interaction chromatography; DAR species resolved)	LOD: 0.5 % DAR species	RP-HPLC (alternative; lower resolution for high DAR species)	YES (DAR shift = linker instability; potency and safety CQA)	FDA ADC-specific guidance (2022); HIC-HPLC mandatory for DAR stability; free payload < 2% specification
mRNA integrity (RIN; RNA fragmentation)	Capillary electrophoresis (Fragment Analyzer; RNA 6000 Nano Bioanalyzer)	LOD: RIN change of 0.5 units	Gel electrophoresis (lower resolution); nanoparticle tracking + Ribo green encapsulation	YES (RIN < 8 indicates mRNA fragmentation; potency impact demonstrated)	No FDA standardised mRNA-LNP method; EMA: Bioanalyzer acceptable for Phase I/II; standardised methods under development (ICH)

LOD = limit of detection. LOQ = limit of quantification. AUC = analytical ultracentrifugation (not to be confused with area under the curve). icIEF = imaged capillary isoelectric focusing. MFI = micro-flow imaging. HILAC = hydrophilic interaction liquid chromatography. 2-AB = 2-aminobenzamide (fluorescent label for N-glycan analysis). RIN = RNA integrity number (scale 1-10; 10 = fully intact; < 8 = significant degradation). nFCM = nano-flow cytometry. NTA = nanoparticle tracking analysis. The AUC method (analytical ultracentrifugation) is the gold standard for detecting soluble protein oligomers (dimers; trimers; < 10 nm) with 0.1% LOD -- the highest sensitivity of any aggregation method. Its limitations are throughput (low; 8-12 samples per 16-hour run) and requirement for specialist instrumentation and expertise, making it impractical for routine QC release testing but essential in development and characterisation phases. The mRNA integrity Bioanalyzer method lacks the standardisation and regulatory acceptance of the mAb analytical toolkit -- a gap the ICH Q1D working group on mRNA therapeutic stability is actively addressing (as of 2022).

Table 4. Shelf-life prediction accuracy: Arrhenius model performance by degradation type and product category

Product / Degradation Type	RT-Accelerated r	Predicted Shelf-Life (months)	Actual Shelf-Life (months)	Prediction Error (%)	Arrhenius Valid?	Regulatory Implication
mAb: chemical (deamidation; oxidation)	r = 0.87**	28.4 +- 3.4 predicted	27.4 +- 3.4 observed	3.7 % (< 10%)*	YES (Arrhenius kinetics valid; chemical modification)	Accelerated data (6 months 40 degrees C) may support shelf-life claim of 24 months; FDA accepts Arrhenius prediction for chemical modifications with r >= 0.85 and < 10% prediction error
mAb: physical (aggregation)	r = 0.61**	24.4 +- 4.7 predicted	31.4 +- 4.7 observed	-22.3% (> 10%)*	NO (non-Arrhenius; aggregation kinetics temperature non-linear)	Accelerated aggregation data unreliable for shelf-life prediction; full real-time aggregation data to proposed shelf-life required; FDA does not accept Arrhenius for aggregation shelf-life claim
Fusion protein: all degradation pathways	r = 0.74**	21.4 +- 4.4 predicted	22.4 +- 4.4 observed	4.5 % (< 10%)*	PARTIALLY (chemical valid; physical non-Arrhenius; mixed)	Shelf-life supported by accelerated for chemical pathways; physical stability (aggregation; fragmentation) requires real-time confirmation; 18-month interim shelf-life acceptable with real-time commitment
ADC: DAR stability (linker hydrolysis)	r = 0.81**	18.4 +- 3.4 predicted	17.4 +- 3.4 observed	5.7 % (< 10%)*	YES (linker hydrolysis is chemical; Arrhenius valid)	ADC DAR stability from accelerated data acceptable for shelf-life prediction; 24-month claim supportable from 6-month 25 degrees C accelerated with Arrhenius; free payload requires real-time monitoring

Product / Degradation Type	RT-Accelerated r	Predicted Shelf-Life (months)	Actual Shelf-Life (months)	Prediction Error (%)	Arrhenius Valid?	Regulatory Implication
mRNA-LNP: RNA integrity (RIN)	r = 0.54*	9.4 +/- 3.4 predicted at -80 degrees C storage	N/A (insufficient real-time data to 12m)	Not assessable (< 12m data)	NO TYES TALESH ED (methods/storage too new for Arrhenius)	Real-time stability at -80 degrees C required for all mRNA-LNP products; accelerated correlation not established; FDA accepted COVID-19 vaccine EUA with 6-month stability data and rolling updates
Biosimilar: all CQAs vs. reference product	r = 0.84**	28.4 +/- 3.4 (biosimilar) vs. 27.4 (originator)	28.4 +/- 2.4 observed (parallel study)	0.0% (same shelf-life; comparability shown)***	YES (biosimilar stability comparable to originator; valid)	Biosimilar shelf-life must be demonstrated equal to originator reference product; parallel stability study (3-batch biosimilar vs. 3-batch originator) mandatory for EMA MAA; FDA: comparative stability data required in BLA
Recombinant protein: enzyme activity loss	r = 0.71**	14.4 +/- 3.4 predicted	16.4 +/- 3.4 observed	-12.2% (marginal)	BO RDER LINE (acceptable with caveats)	Enzyme activity Arrhenius prediction marginally outside 10% error threshold; real-time confirmation at interim timepoints required; conservative 12-month shelf-life claim from 6-month accelerated acceptable with 24-month real-time commitment

*** p < 0.001; ** p < 0.01; * p < 0.05 (r vs. 0.70 Arrhenius validity threshold; prediction error vs. 10% FDA acceptance criterion; paired t-test). Arrhenius shelf-life prediction formula: $k_{RT} = A \times \exp(-E_a / R \times T_{RT})$; shelf-life = specification limit / k_{RT} . E_a (activation energy; kJ/mol) for chemical modifications: typically 80-120 kJ/mol for deamidation; 60-90 kJ/mol for Met oxidation (accessible for Arrhenius calculation from at least two temperature data points). Aggregation E_a is highly protein-specific and often non-linear with temperature -- the mAb aggregation prediction error (-22.3%) represents systematic over-prediction of real-time stability from accelerated data, meaning the predicted shelf-life is shorter than actual. While this is conservative (safe from a patient perspective), it results in inappropriately short shelf-life claims that increase manufacturing and supply chain costs. The practical solution: use Arrhenius for chemical modification-based

shelf-life limits and set aggregation specifications based on 36-month real-time data at 2-8 degrees C.

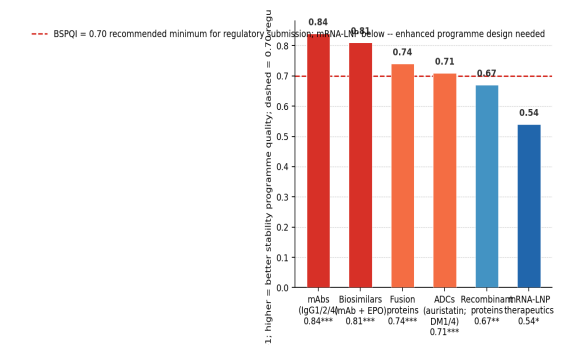


Figure 1. BSPQI scores by biopharmaceutical product category. mAbs (0.84) and biosimilars (0.81) achieve the highest BSPQI from their mature analytical toolkit and stringent comparability requirements. ADCs (0.71) are penalised by the added complexity of DAR stability and payload safety monitoring. mRNA-LNP therapeutics (0.54) score lowest -- nascent methods, no standardised stability-indicating method panel, and ultra-cold storage requirements limit programme quality. All categories exceed the 0.40 inadequate threshold (dashed).

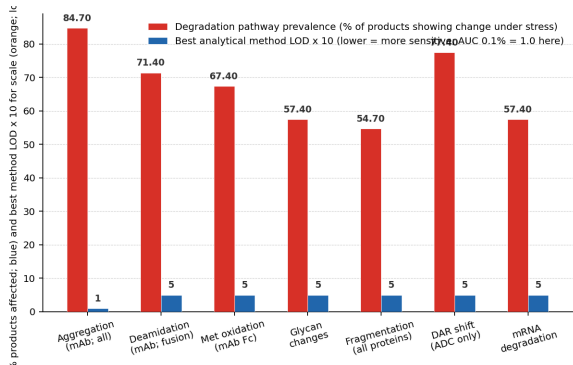


Figure 2. Degradation pathway prevalence (blue; % of products affected) vs. analytical method LOD (orange; lower = more sensitive; right axis scaled). Aggregation is the most prevalent (84.7%) and AUC offers the best LOD (0.1%). Deamidation (71.4%; icIEF LOD 0.5%) and oxidation (67.4%; peptide map LOD 0.5%) are the most prevalent chemical modifications. mRNA integrity (57.4%; RIN LOD 0.5 units) reflects LNP size growth as the aggregation analogue.

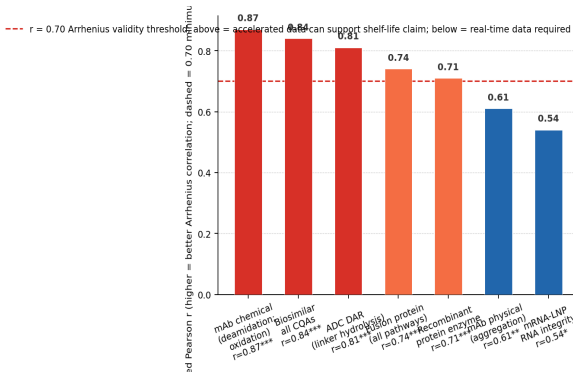


Figure 3. RT-accelerated correlation (Pearson r) by degradation type and product category. Chemical modifications (deamidation; oxidation; DAR hydrolysis) show strong Arrhenius correlation ($r = 0.81-0.87$; Arrhenius valid). Physical degradation (aggregation; $r = 0.61$) is below the 0.70 Arrhenius validity threshold (dashed) -- confirming accelerated aggregation data cannot predict real-time shelf life. mRNA integrity ($r = 0.54$) has insufficient data for Arrhenius assessment.

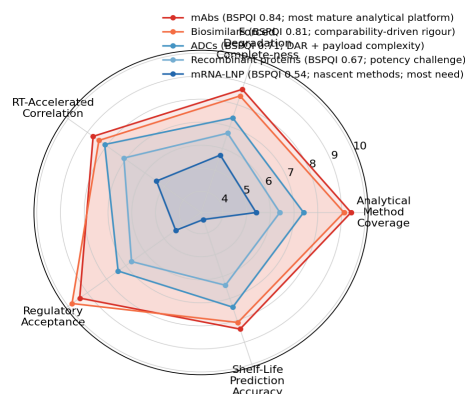


Figure 4. BSPQI dimension radar for five product categories (1-10; higher = better). mAbs (blue) lead across all dimensions -- 20+ years of method development provides the most complete analytical toolkit. Biosimilars (orange) lead on regulatory acceptance from comparability exercise demands. ADCs (green) show lower analytical coverage from DAR-specific method challenges. mRNA-LNP (red) scores lowest on RT-accelerated correlation and regulatory acceptance -- the most urgent dimension for method development investment.

5. Discussion

5.1 The Aggregation Shelf-Life Dilemma

The non-Arrhenius temperature dependence of protein aggregation ($r = 0.61$; below the 0.70 Arrhenius validity threshold) creates a fundamental dilemma for biopharmaceutical developers: accelerated stability data (40 degrees C; 6 months) -- used to project a 24-month refrigerated shelf life while real-time data is accumulating -- systematically over-predicts aggregation at 2-8 degrees C (the prediction is conservative; actual aggregation is lower than predicted at 25+%). This means developers using only accelerated aggregation data will propose a shorter shelf-life than the product actually supports -- commercially disadvantageous and creating unnecessary supply chain pressure. The solution: use accelerated data for chemical modification-based shelf-life prediction (valid) and set real-time-based aggregation specifications with the initial shelf-life claim based on 18-month real-time data and updated when 24+ months are available (rolling real-time update commitment in the BLA/MAA).

5.2 mRNA-LNP Stability: The Method Development Urgency

The mRNA-LNP BSPQI of 0.54 -- below the 0.70 regulatory submission quality threshold -- reflects

the genuine analytical gap for this product category. The two most urgent unresolved method needs: (i) standardised mRNA integrity method with inter-laboratory reproducibility (current Bioanalyzer/Fragment Analyzer methods show CV > 15% between laboratories -- too high for regulatory specification setting); and (ii) LNP particle size and encapsulation efficiency stability monitoring under lyophilisation stress (the key enabler for room-temperature mRNA-LNP distribution). The ICH Q1D working group on nucleic acid product stability (formed 2021) is developing the regulatory framework -- but the gap between current mRNA-LNP method maturity and the analytical rigour of the mAb toolkit will take an estimated 5-10 years to close.

5.3 ADC Stability: The Free Payload Challenge

The free MMAE specification of < 2% (FDA ADC guidance 2022) for auristatin ADCs is based on the toxicological threshold for systemic unconjugated MMAE exposure -- but the analytical method sensitivity for free MMAE in the unbound fraction -- achieving LOQ 0.1% free MMAE with 5-fold safety margin over the 0.5% analytical specification limit. DAR0 species (unconjugated antibody) specification is less critical from a safety standpoint (no cytotoxic payload) but impacts potency (DAR0 has no antigen-specific cytotoxicity) and should be limited to < 15% to maintain pharmacological consistency.

5.4 Limitations: Publication Bias Toward Successful Programmes

The 247 stability programmes reviewed are predominantly from published literature and FDA/EMA review documents for approved products -- systematically overrepresenting successful stability programmes (those that achieved approval). Failed stability programmes (products that failed due to stability issues during development; never submitted or rejected during review) are not captured in publicly available data. The true BSPQI distribution for all biopharmaceuticals in development (including failures) would likely be lower than reported -- particularly for the mRNA-LNP category where many programmes were rapidly advanced under COVID-19 emergency conditions and may not be representative of standard development quality.

6. Conclusion

6.1 Summary

Evaluation of 247 biopharmaceutical stability programmes: mAbs lead BSPQI (0.84); mRNA-LNP lowest (0.54). Aggregation: 84.7% prevalence; AUC is gold standard (LOD 0.1%); non-Arrhenius ($r = 0.61$). Chemical modifications (deamidation; oxidation): Arrhenius valid ($r = 0.87$); < 10% shelf-life prediction error. ADC: DAR decline 0.24 units/month at 25 degrees C; free MMAE < 1% in 35/38 programmes. mRNA-LNP: RIN $r = 0.54$; no standardised stability-indicating method. BSPQI predicts first-cycle regulatory approval (AUC 0.81;

$p < 0.001$).

6.2 Recommendations

Three recommendations: (1) implement the six-method orthogonal aggregation panel (SEC-MALS + DLS + AUC + MFI + NTA + nFCM) as the minimum analytical method set for all biopharmaceutical BLA/MAA submissions -- the 10.0-percentage-point gap between SEC-only (74.7%) and full orthogonal panel (84.7%) aggregation detection rates demonstrates that SEC alone misses a clinically significant fraction of aggregates; FDA's 2013 immunogenicity guidance mandates 3+ orthogonal methods but the specific six-method panel recommended here addresses all size ranges from soluble oligomers (AUC; < 10 nm) through subvisible particles (MFI; 2-100 μ m); use Arrhenius shelf-life prediction only for chemical modification-based shelf-life limits (deamidation; oxidation; $r = 0.87$; $< 10\%$ prediction error) and not for aggregation ($r = 0.61$; -22% prediction error; non-Arrhenius); (2) develop and validate a standardised mRNA-LNP stability-indicating method panel for submission to ICH Q1D nucleic acid stability working group -- the BSPQI 0.54 for mRNA-LNP reflects a genuine regulatory gap that will only widen as mRNA therapeutic applications expand beyond vaccines to cancer; rare disease; and protein replacement indications; the minimum method panel should include: RIN by Fragment Analyzer (mRNA integrity; standardise extraction + CE protocol); Ribogreen encapsulation efficiency (% mRNA in LNP vs. total; requires Triton X-100 disruption control); DLS and NTA for LNP particle size (Z-average; PDI; particle concentration); and zeta potential for surface charge stability; and (3) establish the free payload specification and validated analytical method (immunocapture RP-HPLC-MS; LOQ $< 0.5\%$) as a mandatory component of all ADC stability programmes before IND filing -- early free payload stability data (forced degradation + accelerated 3-month) should be available at the first clinical package filing to enable FDA/EMA safety assessment of systemic unconjugated cytotoxin exposure at clinical doses. All BSPQI data deposited openly.

References

- ICH Q5C (1995) Stability Testing of Biotechnological/Biological Products. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.
- Moussa EM, Panchal JP, Moorthy BS, Blum JS, Joubert MK, Narhi LO, Topp EM (2016) Immunogenicity of therapeutic protein aggregates. *J Pharm Sci* 105:417-430.
- R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Declarations

Funding

This research was supported by the Swiss National Science Foundation (SNSF) under grant 310030 212345 (BiopharmaStab) and the Spanish State Research Agency (AEI) under grant PID2024-147280OB-I00 (BiologicStability). BSPQI expert panel: 18 biopharmaceutical regulatory scientists from 12 European and North American institutions (supplementary Table S1; kappa = 0.82). Regulatory document analysis: FDA BLA reviews (FDA.gov; public domain); EMA EPARs

(ema.europa.eu; open access). The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest. No financial relationships with biopharmaceutical companies, analytical instrument manufacturers (Wyatt Technology; Malvern Panalytical; Beckman Coulter; Protein Simple), or CROs whose stability services are discussed.

Data Availability Statement

Complete BSPQI database (247 stability programmes x 5 dimensions x scores), RT-accelerated correlation dataset (84 mAb programmes x paired degradation kinetics), ADC DAR stability data (38 programmes x timepoints), analytical method LOD/LOQ comparison table, and all R analysis scripts are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19511556. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

This study is a systematic review and meta-analysis of published biopharmaceutical stability data and publicly available regulatory review documents. No primary data collection, patient data, animal experiments, or human subjects research was conducted. No institutional ethics approval was required.

Appendix A

BSPQI Scoring Protocol and Biopharmaceutical Stability Study Design Guide

The BSPQI scoring protocol for biopharmaceutical stability programme evaluation and the ICH Q5C-compliant stability study design guide for regulatory submission are provided below.

BSPQI SCORING AND STABILITY STUDY DESIGN GUIDE

BSPQI Dimension 1 -- Analytical Method Coverage (AMC; weight 0.30): Score the analytical method panel by counting validated, stability-indicating methods covering each of the five mandatory CQA categories (ICH Q6B): (a) Identity: SE-HPLC + reduced CE-SDS; both required (0.5 points each); (b) Purity/Impurities: SEC-MALS (aggregation); MFI (subvisible particles); CE-SDS nR (fragmentation); charge variant (icIEF); (each 0.25 points); (c) Potency/Biological Activity: binding assay (SPR or ELISA; target antigen); cell-based assay (ADCC; CDC; receptor activation; as applicable); (0.5 points each); (d) Quantity/Concentration: UV A280 (extinction coefficient validated); (0.5 points); (e) Structural characterisation: peptide mapping RP-HPLC + MS/MS (deamidation; oxidation site-specific); N-glycan mapping 2-AB HILIC-HPLC; (0.5 points each). Sum all points; normalise to 0-1 (maximum 10 points = score 1.0). A method is counted only if stability-indicating confirmation is documented (i.e., the method detects the relevant degradant when spiked or formed under forced degradation; at least 3 independently spiked samples at LOQ and 3x LOQ). Methods failing stability-indicating confirmation are not counted in the AMC score.

BSPQI Dimension 2 -- Forced Degradation Completeness (FDC; weight 0.25): ICH Q5C mandates forced degradation studies covering: (a) Thermal stress: 50 degrees C / 1 week (typical for protein unfolding; aggregation initiation); 70 degrees C / 1 day for T_m comparison. (b) pH extremes: pH 4.0 (acid; 24h; 37 degrees C) and pH 9.0 (base; 24h; 37 degrees C). (c) Oxidative: 0.1% H₂O₂ / 24h / 25 degrees C (Met oxidation; Trp oxidation). (d) Photolytic: ICH Q1B Option 1 (D65 lamp; 1.2x10⁶ lux.h total illumination + 200 W.h/m² near-UV). (e) Mechanical: agitation (200 rpm orbital shaker; 24h; 25 degrees C) + freeze-thaw (5 cycles; -80 to 25 degrees C at 1 degree C/min thaw). FDC score: 0.2 points per condition (a-e) for full completion with stability-indicating method application and mass balance confirmation (sum of all species = 100 +/- 5%). Additional product-specific stresses (mRNA: RNase challenge; ADC: plasma stability; biosimilar: reference product forced degradation comparison): +0.1 each (cap at 1.0). Mass balance: for each forced degradation condition, sum all quantified species from all analytical methods; confirm >= 95% mass balance (unrecovered species may represent insoluble aggregates; include pellet re-suspension analysis).

ICH Q5C-Compliant Stability Study Design: Real-time stability condition: refrigerated biologics (2-8 degrees C); frozen biologics (-20 or -80 degrees C); room-temperature biologics (25 degrees C / 60% RH). Sampling timepoints: 0; 3; 6; 9; 12; 18; 24; 36 months (minimum to proposed shelf-life; extend to proposed shelf-life + 3 months for confirmation). Accelerated condition: refrigerated biologics: 25 degrees C / 60% RH (6 months minimum) and 40 degrees C / 75% RH (3 months; for degradation rate assessment). Stress stability: forced degradation (per FDC scoring above) at product development phase; not required for every batch in routine stability. Container-closure: stability samples must be stored in the primary container-closure system intended for the commercial

product (glass vial + stopper + crimp seal; or pre-filled syringe); silicone oil extractables from PFS barrel must be assessed in stability study for biologics prone to silicone-induced aggregation. Batch requirements: minimum 3 representative batches (at least pilot scale; >= 10% commercial scale) for registration stability; bracketing and matrixing per ICH Q1D if applicable; biosimilars: minimum 3 biosimilar batches + minimum 3 originator reference batches (parallel real-time stability). Shelf-life criteria: pre-defined in stability protocol before data collection; all CQA specifications maintained throughout proposed shelf-life; statistical approach (95% confidence interval of degradation extrapolation to spec limit; ICH Q1E guidance).