


REVIEW

Adverse immunostimulation caused by impurities: The dark side of biopharmaceuticals

Joannes A. A. Reijers¹  | Karen E. Malone² | Jeffrey J. Bajramovic³ | Richard Verbeek⁴ | Jacobus Burggraaf¹ | Matthijs Moerland¹

¹Centre for Human Drug Research, Leiden, the Netherlands

²Sagacity, Leiden, the Netherlands

³Alternatives Unit, Biomedical Primate Research Centre, Rijswijk, the Netherlands

⁴Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, Leiden, the Netherlands

Correspondence

Matthijs Moerland, Centre for Human Drug Research, Leiden, The Netherlands.
Email: mmoerland@chdr.nl

Drug safety is an important issue, especially in the experimental phases of development. Adverse immunostimulation (AI) is sometimes encountered following treatment with biopharmaceuticals, which can be life-threatening if it results in a severe systemic inflammatory reaction. Biopharmaceuticals that unexpectedly induce an inflammatory response still enter the clinic, even while meeting all regulatory requirements. Impurities (of microbial origin) in biopharmaceuticals are an often-overlooked cause of AI. This demonstrates that the current guidelines for quality control and safety pharmacology testing are not flawless. Here, based on two case examples, several shortcomings of the guidelines are discussed. The most important of these are the lack of sensitivity for impurities, lack of testing for pyrogens other than endotoxin, and the use of insensitive animal species and biomarkers in preclinical investigations. Moreover, testing for the immunotoxicity of biopharmaceuticals is explicitly *not* recommended by the international guidelines. Publication of cases of AI is pivotal, both to increase awareness and to facilitate scientific discussions on how to prevent AI in the future.

KEYWORDS

adverse drug reactions, biopharmaceutics, drug safety, toxicology

1 | INTRODUCTION

Whenever a new pharmaceutical enters the clinic, it is subjected to rigorous testing, to ensure the safety of patients. Guidelines and regulations are designed to maintain the high standard expected by the public. Agencies and inspectors oversee the correct implementation of these regulations and guidelines by manufacturers and researchers. This is especially important when humans are exposed for the first time to a new drug product.

In 2006, TGN1412, a T cell superagonistic anti-CD28 monoclonal antibody, was administered to six healthy volunteers. Following drug

administration, the trial participants developed multiorgan failure resulting from what was later termed a “cytokine storm”.¹

This case underlined the potentially lethal consequences of a systemic inflammatory reaction induced by a pharmaceutical, and thus the importance of studying a pharmaceutical's immunostimulatory propensity during preclinical development. One would expect, therefore, that part of the current preclinical testing strategy is directed toward detecting activation of the immune system. However, we know from personal experience that unregistered drug products, mostly biopharmaceuticals, can unexpectedly cause adverse immunostimulation (AI; see Box 1) when administered to human

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2019 The Authors. British Journal of Clinical Pharmacology published by John Wiley & Sons Ltd on behalf of British Pharmacological Society.

Box 1 Adverse Immunostimulation

Immunotoxicity caused by pharmaceuticals has been classified by the Food and Drug Administration (FDA) into five groups,² which partly overlap with the pathophysiologically orientated categories α - γ of immune-related adverse effects described by Pichler³ (Table 1). Pichler proposed a classification specifically for biopharmaceuticals, which also categorized non-immunotoxic adverse reactions (types δ and ϵ). Regarding immunopathological phenomena, type α reactions (termed “adverse immunostimulation” [AI] by the FDA) and hypersensitivity (type β), particularly anaphylaxis, are the most dangerous, for they can quickly become lethal. Anaphylaxis and type α reactions have in common that unbalanced propagation of an immune reaction occurs with systemic release of various mediators, which ultimately leads to widespread organ dysfunction.⁴⁻⁷ In type α reactions, these mediators are cytokines,³ resulting in a so-called cytokine storm.^{6,7} The source of these cytokines is the administration of (high doses of) cytokines or the release of cytokines following treatment with a biopharmaceutical.³ This last category includes TGN1412, muromonab-CD3 and other monoclonal antibodies.⁸⁻¹¹ Another, often overlooked, cause for cytokine release is the presence of impurities or contaminants within pharmaceuticals, typically of microbial origin, triggering an immune response.

It should be noted that a subtle difference exists between the original definition of type α reactions and AI: in AI, high cytokine levels are not necessarily part of the pathogenesis, and AI can also include chronic inflammation.² Later modifications to Pichler's classification^{10,12} broadened the definition of type α reactions toward a general proinflammatory concept, thus becoming more or less synonymous with the FDA's definition of AI. To complicate uniformity further, other terms appear in literature—for example, (acute) infusion reactions and flu-like syndromes, which may refer to type α or type β reactions, or to both.^{10,12-14} For the sake of simplicity, the term “AI” is used throughout this article to describe acute, systemic, inflammatory reactions. It can be a matter of discussion whether these reactions may be classified as classical type α reactions.

volunteers, even while meeting all criteria set forth by the applicable guidelines (see Box 2 and Table 2).

At the Centre for Human Drug Research (CHDR), 10–15 first-in-human clinical trials are executed yearly, of which approximately 40% involve biopharmaceuticals. In the past 5 years, we have

TABLE 1 Classification of immunopathological phenomena caused by biopharmaceuticals

Pathophysiological classification ³	Immunotoxicity classification ²
Type α	Adverse immunostimulation
High cytokine levels	
<ul style="list-style-type: none"> • cytokine administration (augmented primary pharmacology) • cytokine release 	
Type β	
Immunogenicity (immune response against pharmaceutical)	
<ul style="list-style-type: none"> • hypersensitivity (drug allergy), including type IV (T cell mediated) • neutralizing or non-neutralizing anti-drug antibodies 	Hypersensitivity Immunogenicity
Type γ	
Immune deviation (by pharmaceutical)	
<ul style="list-style-type: none"> • immunosuppression, immunodeficiency • immune imbalance or enhancement ○ autoimmunity ○ exacerbation of existing atopy ○ induction of atopy 	Immunosuppression Autoimmunity

The pathophysiological classification³ includes non-immunotoxic adverse reactions—namely, type δ (cross-reactivity) and type ϵ (non-immunological effects). Type α reactions are considered as adverse immunostimulation by the Food and Drug Administration,² but the terms are not synonymous as adverse immunostimulation can also include chronic inflammation, and the underlying mechanism does not require high cytokine levels.

Box 2 Testing for Adverse Immunostimulation (AI)

Before a new drug product can be studied clinically, sufficient evidence needs to be supplied regarding its safety. Tests performed to substantiate these claims fall into two main categories: quality control and safety pharmacology. Quality control encompasses laboratory assessments to verify the compound's identity by structure, amino-acid sequence, physical and chemical properties, and receptor affinity and potency, and includes assessments for (product-related) impurities and contaminants. Preclinical pharmacology studies are directed toward identifying safety issues as a result of exaggerated intended pharmacodynamics (primary pharmacology) or unintended toxic effects. These studies are usually a mixture of *in vitro* (animal/human cell lines) and *in vivo* (animal) experiments. Clinical experiments with investigational medicinal products (IMPs) are executed once sufficient evidence has been gathered during the preclinical phases of development that the product is safe. A summary of the commonly applied tests to detect AI by IMPs is provided in Table 2.

First step: Quality control

All parenterally administered pharmaceuticals have to conform to the requirements regarding sterility and endotoxin content, which are captured in the national, yet harmonized, pharmacopoeias. Sterility is considered proven by a negative culture result for a predefined fraction of the produced batch, usually after 14 days in a suitable medium. For endotoxin, a margin of 5.0 IU per kg body weight per hour is internationally accepted (Table 2). The limulus amoebocyte lysate assay is the preferred method for measuring endotoxin levels and has largely replaced the rabbit pyrogen test.

Biopharmaceuticals, such as X and Y, should be tested specifically for the presence of impurities. Residual cellular components of the manufacturing platform are recognized as potential triggers of the immune system, usually residual DNA and/or host cell proteins. Regulatory guidelines do not prescribe specific assays and acceptance criteria, but introduce a general concept of using validated, appropriate methods and setting (strict) upper limits, which need to be justified.¹⁵ The exact battery of tests and criteria are reviewed and approved on a case-by-case basis by the regulatory authorities. Upon approval, these quality control measures are applied as release specification for individual batches of the drug product.

Second step: Safety pharmacology

For chemically derived compounds, an initial screen for potential immunotoxicity is mandatory, but this is not the case for biopharmaceuticals because they are target specific by design.^{16,17} Indications for potential immunotoxicity are derived from animal experiments. Relevant *in vivo* signs to detect AI include changes in (differential) leucocyte counts and globulin levels. Other useful parameters in assessing immune stimulation are part of the standard safety battery and include body temperature, heart rate and respiratory rate (Table 2).

Third step: Clinical pharmacology

Specific tests for immunotoxicity in humans are not required, unless there is an indication that the drug candidate is potentially immunotoxic. Testing for immunogenicity by biopharmaceuticals is usually included because it is a known problem of this drug class.¹⁸ Standard safety markers, such as haematology, clinical biochemistry, vital signs, as well as reports of adverse events, can be an indication of AI, provided that they are measured sufficiently frequently. Important factors in the trial design to improve the safety of study participants include dose selection, entry criteria (eg, healthy volunteers vs patients), safety window and data monitoring committee.^{19,20}

encountered two biopharmaceuticals that unexpectedly induced AI. In this review, we share our experience with these products, illustrating how the currently employed testing strategy failed in detecting AI. Despite strict adherence to the pertaining guidelines, as well as careful scientific and ethical review, AI was only first detected when these biopharmaceuticals were administered to humans. These cases are not, or only partly, in the public domain and will be denoted X and Y, respectively.

2 | CASE X

X is a recombinant human apolipoprotein A-I Milano, with a molecular weight of roughly 55 kDa, which was developed under the code name ETC-216.²⁸ It is expressed in the periplasmic space of *Escherichia coli*, and was manufactured at fully certified and accredited facilities in the US and Western Europe. Only compendial (pharmacopoeia grade) materials were used, and when these were not available, the materials used had to adhere to predefined internal standards. Sterility was confirmed by the membrane filtration method (*European Pharmacopoeia*, Ph.Eur.,²⁵ paragraph 2.6.1; *US Pharmacopoeia*, USP,²⁶ general chapter <71>) and all release criteria were met in terms of physical appearance, concentration and desired activity. Details on the impurity content and endotoxin levels are presented in Table 3.

The single-dose toxicity of the drug substance was studied in rats and cynomolgus monkeys, up to a dose more than 10 times the highest dose ever administered to humans (allometrically scaled). Repeated-dose toxicity studies were performed with the same dose levels in rats for up to 2 weeks, and in monkeys for up to 6 weeks. Only effects that could be attributed to exaggerated pharmacology were observed, and those were completely reversible. No changes in behaviour, haematology or clinical chemistry occurred, and there were no relevant findings at autopsy.

Additionally, the pharmacodynamic activity of the drug, albeit from a different lot, had been studied in rabbits and in a mouse model of atherosclerosis. In this research, multiple animals died after repeated doses. The cause of these fatalities in mice was judged to be the induced cardiovascular disease. In rabbits, anaphylaxis developed, which was considered to be a response to repetitive exposure to a foreign protein.

In an unpublished phase I single-dose study in healthy volunteers ($n = 28$, five dose levels), X induced dose-dependent increases in the number of neutrophils, with a maximum at 4 hours post-administration. In addition, the incidence of gastrointestinal symptoms – occurring 2–4 hours after infusion – rose with each dose escalation, as did the incidence of diaphoresis and fever.²⁹

In the next phase, in patients with an acute coronary syndrome,³⁰ one patient out of a group of 22 receiving the highest dose level experienced what was described as a hypersensitivity reaction. A second (unpublished) patient trial was quickly suspended after one patient on active treatment developed a severe reaction during infusion, leading to multiorgan failure.

TABLE 2 Overview of current testing strategy

Category/parameter	Method	Acceptance criterion	Guidelines ^a
Quality control	Laboratory assessments		Q5D
Sterility (bioburden)	Culture	No growth	Ph.Eur. 2.6.1, USP <71>
Endotoxin (LPS)	LAL test	<5.0 IU/kg/h (body weight) ^b	Ph.Eur. 2.6.14, USP <85>
	(rabbit) pyrogen test	<0.5 °C increase in 3 h ^c	Ph.Eur. 2.6.8, USP <151>
	Monocyte activation test	Reaction less than the reaction induced by the allowed endotoxin contamination ^b	Ph.Eur. 2.6.30
Host cell impurities^d	Identified and quantitated		
Residual DNA	PCR/hybridization	<10 ng/dose ^e /(strict) upper limits, as appropriate	Ph.Eur. 01/2008:0784, USP <1045> ^{,21}
Proteins (HCPs)	Immunoassay (ELISA), Western blot	(strict) upper limits, as appropriate	Ph.Eur. 01/2008:0784, USP <1045> ^{,21}
Virus ^f		Maximal clearance	Q5A, Ph.Eur. 01/2008:0784, USP <1045>
Safety pharmacology^g	Animal toxicology		S6
Safety battery			
Central nervous system toxicity	Body temperature	Evaluation and interpretation	S7A
Cardiovascular toxicity	Heart rate		
Respiratory toxicity	Respiratory rate		
Immunotoxicity^h			S8, ²
Standard	(differential) leucocyte count		
	Globulin levels		
Additional ⁱ	Immune function (e.g. T cell-dependent antibody response, natural killer cell activity, host resistance, cell-mediated immunity)		
Clinical pharmacology^g	Human toxicology		

Abbreviations: ELISA, enzyme-linked immunosorbent assay; HCPs, host cell proteins; LAL, limulus amoebocyte lysate; LPS, lipopolysaccharide (endotoxin); PCR, polymerase chain reaction.

^aGuidelines include *ICH Harmonised Tripartite Guidelines* (Q5A,²² Q5D,²³ S6,¹⁶ S7A,²⁴ S8¹⁷), *European Pharmacopoeia* general texts and monographs (Ph. Eur.²⁵), *US Pharmacopoeia* general chapters (USP²⁶), and other guidelines or directives (denoted by their reference number).

^bNotwithstanding pharmacopoeia monographs, a margin of 5.0 IU (or endotoxin unit, EU) per kg body weight per hour, or a dosing equivalent, is usually considered acceptable for parenterally administered products, with the exception of radiopharmaceuticals and intrathecally administered products, for which a lower limit is set.

^cRespective European and US pharmacopoeias are not harmonized; they differ in the number of animals included in a retest, whether maximal individual temperature responses or summed responses are used, and the exact acceptance criteria.

^dOnly recommended for biopharmaceuticals.

^eA limit of 10 ng per dose is commonly suggested,²⁷ although it is not absolute.^{2,5}

^fViral inactivation or removal is usually only tested on the cell line, not on the drug substance.

^gOnly required for marketing authorization, not for control of different batches (release specification).

^hA limit of 10 ng per dose is commonly suggested.²⁷

ⁱOptional.

Later investigations demonstrated elevations in circulating interleukin (IL) 6 and tumour necrosis factor- α (TNF- α) levels following administration of X in humans, which were traced back to several host cell proteins (HCPs) within X, one of which was flagellin.^{31,32} The

manufacturer required more than a decade to reduce or eliminate these HCPs, but in the end X could be successfully reintroduced. Clinical studies with the new drug product confirmed the absence of any cytokine response both in healthy volunteers and patients.^{33,34}

TABLE 3 Product specifications

Structure	X	Y
	Polypeptide dimer	Polypeptide monomer
Molecular weight (kDa)	~55	~45
Expression method	<i>Escherichia coli</i>	<i>Escherichia coli</i>
Sterility (membrane filtration)	Sterile	Sterile
Endotoxin ^a (LAL assay) (IU/kg/h)	<2.3	<0.4
Residual DNA ^a (ng)	3 ^b	0.3 ^c
HCPs (ppm)	≤10	≤50
Aggregates (SE-HPLC) (%)	<0.6	Not detected

Abbreviations; HCPs, host cell proteins; LAL, limulus amoebocyte lysate; SE-HPLC, size exclusion high-performance liquid chromatography.

^aResults are based on the highest (single) dose administered to humans.

^bThreshold technique.

^cQuantitative polymerase chain reaction.

3 | CASE Y

Y is a mutant form of a human plasma protein, approximately 45 kDa in size. Several years earlier, the wild-type protein had been manufactured via recombinant DNA techniques, and tested extensively in dozens of clinical trials. Its track record was spotless in terms of safety, other than the risks associated with its intended pharmacological action. In particular, there was no indication of AI in any of the performed trials, and nor was it anticipated, based on the mode of action. The missense mutation (Y) was produced to increase the stability of the protein, and this product had been tested in many animal models (rats, dogs, rhesus monkeys) before clinical development commenced. None of the animals ever demonstrated the signs or symptoms of an inflammatory reaction, even at toxic doses.

E. coli was used to produce pharmacological grade material of Y at a fully certified and accredited facility in Western Europe. All raw materials conformed to US or EU pharmacopoeias, or were controlled by internal specifications. Sterility was confirmed by the membrane filtration method (Ph.Eur.²⁵ 2.6.1, USP²⁶ <71>) and all release criteria were met, in terms of physical appearance, concentration and desired activity. Details on the impurity content and endotoxin levels are presented in Table 3.

Toxicity was examined in rats and cynomolgus monkeys at a dose more than 10 times the highest dose ever administered to humans (allometrically scaled). Only single-dose toxicity studies were performed because the product was being developed as a one-dose treatment. No effects other than those expected to result from exaggerated pharmacology were detected.

In humans, dose-dependent increases in the incidence of gastrointestinal complaints and fever were observed following treatment with Y (single dose, $n \approx 10$). The severity of the reaction correlated with the peak value for C-reactive protein and plasma cytokine levels (IL-6 and TNF- α). Subsequent analyses into the cause of the events revealed that Y caused toll-like receptor (TLR) 4-mediated activation,

despite a negative limulus amoebocyte lysate (LAL) test. In addition, using a different method, >10% protein aggregates were detected in Y. It took more than 3 years to produce a new pharmacological-grade product with reduced levels of endotoxin, protein aggregates, residual DNA (rDNA) and host cell proteins. When the latter product was administered to healthy volunteers, no AI occurred at much higher plasma levels than were previously reached.

4 | THE WEAKEST LINK

These two cases raise the questions of how the causative impurities remained undetected, and why their immunostimulatory propensity was not recognized at an earlier stage, especially as the development was not marked by carelessness. On the contrary, the products were being manufactured by renowned companies, and studied and tested by dedicated scientists. The guidelines were meticulously followed, and more stringent than necessary acceptance criteria were applied (Tables 2 and 3). Yet, despite this, unforeseen AI occurred in the clinical studies. In the following paragraphs, we identify the most important shortcomings of the current testing strategy which allowed the culprits of AI in X and Y to escape detection in the preclinical phases of development.

4.1 | It's all in the number: Test sensitivity

Quality control is essentially based on using specific, validated methods for detecting unwanted components in pharmaceuticals. However, all laboratory tests have their limitations and, in particular, insufficient sensitivity can be problematic. For example, Y met the endotoxin specification when tested with the LAL assay (Tables 2 and 3); yet, subsequent investigations demonstrated TLR4-mediated activation, and a specific endotoxin enzyme-linked immunosorbent assay (ELISA) indicated that actual levels were 10–20-fold higher than previously measured.

Although the definitive cause of the negative result in the LAL assay remains to be elucidated, it is known from the literature that the LAL assay cannot detect certain endotoxins, such as low-molecular-weight endotoxin.³⁵ The tests for other impurities can also give spuriously low results. Most commercially available assays for rDNA utilize an amplification technique which allows detection of only a selection of rDNA.³⁶ Likewise, assays for HCP quantification, typically ELISAs, measure only proteins against which antibodies were raised during the development of the assay, and do not capture other sources of contamination.^{37,38} The latter was the case with X, in which a process-specific ELISA measured an HCP concentration that was a factor of 3–4 higher than the initial result using a commercial, generic *E. coli* HCP ELISA.

4.2 | Behind the number: Potency

Specific, quantitative tests are routinely applied in quality control settings, as clear limits of acceptance can be defined. Although assays for

tested impurities and contaminants *seemingly* provide black and white cut-offs, they are limited in predicting *in vivo* toxicity. For example, the LAL assay yields no information on the biological potency of a given endotoxin, which can differ between bacterial strains by a factor of 10,000.^{39,40} Additionally, a strict cut-off for one impurity does not take into account the potentially synergistic effects of multiple impurities. This is an important limitation as it is conceivable that multiple impurities are copurified at any given time during the production process. Both for X³² and Y, it was not a single HCP or other bacterial product that could be identified as cause for the AI; it was the reduction in the total load of nonhuman material other than the pharmacologically active substance which resulted in the absence of AI in trial volunteers upon reintroduction into the clinic.

4.3 | Number and species: Animal toxicology

AI resulting from impurities can remain undetected in safety pharmacology studies as the relevant immunological pathway may not exist in the particular animal species or is different to that in humans. In addition, the sensitivity for bacterial products, such as endotoxin, can differ.⁴¹ In particular, the relative lack of comparability between human and murine immune systems is noteworthy.^{42,43} However, the guidelines also call for short-term toxicity studies in a nonrodent species, in case a pharmacologically relevant nonrodent species exists.¹⁶ Commonly, a primate is used for such studies with biopharmaceuticals.⁴⁴

Although the aforementioned factors may contribute to the failure to recognize AI, a subsequent investigation with X demonstrated that it elicited a cytokine response in cynomolgus monkeys similar to that in humans, with all the associated clinical features, such as fever, tachycardia and an increase in the white blood cell count. The real issue in this case therefore seems to be that either appropriately sensitive measures of AI (eg, circulating cytokines) had not been selected initially or that measurements had been taken too infrequently to detect AI (eg, vital parameters, haematology results).

4.4 | Blind spot

Surprisingly, testing for immunopathological effects is not included in the “Safety Pharmacology Core Battery”, and in addition “routine testing” is explicitly *not* recommended by the International Conference on Harmonisation guidelines for biopharmaceuticals,^{16,24} even though biopharmaceuticals have an inherent risk of AI caused by impurities, especially if manufactured in bacterial cell systems. Testing for toxicities is not required if they are not reasonably expected to occur.¹⁵ For biopharmaceuticals, for example, when there is extensive knowledge and data on receptor distribution and function, and the biopharmaceutical under investigation has a selectively high affinity for that receptor, dedicated safety pharmacology studies may be eliminated entirely or in part. For biopharmaceuticals intended for end-stage cancer treatment, even further exceptions to the requirements are possible.^{24,45} This guidance may result in minimal safety testing of a compound before entering the clinic.

Even when dedicated immunotoxicity studies are performed, the emphasis seems to be on long-term immunosuppression or enhancement by evaluating macroscopic pathology, organ weights and histology.^{2,17} This is at odds with the observation that undesired inflammatory reactions are commonly transient and rarely manifest themselves in histological changes. Case X highlights how the AI can be overlooked in safety pharmacology experiments as well as during clinical trials. AI in clinical trials may be easily misinterpreted as hypersensitivity or be grouped under a nonspecific term such as “infusion reaction”.

5 | LESSONS FROM X AND Y

In the aftermath of the TGN1412 tragedy, an Expert Scientific Group made 22 recommendations to increase the safety of participants in (first-in-human) clinical trials.^{46,47} Many of these were reiterated following the more recent BIA 10-2474 trial disaster.^{48,49} The proposed risk assessments^{46,48,50} focus on establishing the likelihood of unanticipated adverse effects of new drug substances entering the clinic, especially for those with complex and novel mechanisms. When applying this strategy to compounds X and Y, and specifically to (the novelty of) the mode of action and knowledge about human exposure to similar substances, it must be concluded that both X and Y were correctly labelled as safe.

The real threat to safety was the uncharacterized components in X and Y, impurities, the presence of which was not disputed (Table 3), but the potential risk involved was underestimated. This line of reasoning is common, as is also apparent from the importance placed on dedicated immune toxicity studies in the international guidelines. These are explicitly *not* recommended for biopharmaceuticals.¹⁶ Furthermore, the immune system is regarded as being “of less immediate investigative concern” because it is an organ system, “the functions of which can be transiently disrupted by adverse pharmacodynamic effects without causing irreversible harm”.²⁴

The outcome of the risk assessment for the uncharacterized components should immediately raise several red flags. There is usually little knowledge about the effects of impurities in humans or animals, and safe levels are often not available. Multiple signalling pathways may be triggered and the impurity can act via the immune system. All of these factors should have placed trials with X and Y in a high-risk category,⁴⁸ and, considering the bacterial origin of the impurities, an immune-mediated effect could have been anticipated. The lack of emphasis on immunotoxicity caused by biopharmaceuticals is therefore striking.

It is important to investigate and reduce the adverse immunostimulatory propensity of biopharmaceuticals for the safety of trial volunteers and future patients, but it is also desirable, from a business perspective, to identify and possibly eliminate such characteristics in an early stage of development. Each step forward increases the expense of going back to the drawing board, as well as the financial loss associated with ultimate failure and drug withdrawal. Discovering severe AI during clinical studies usually results in delays of many

years, cases X and Y being no exception. It may even lead to withdrawing the drug from a company's development pipeline, no matter how promising the compound.

Cases X and Y did not make the news, or lead to a public outcry, making the need for changes less pressing, even though details of the clinical trials with X and Y are known to regulatory agencies and health authorities around the globe. The lack of publicly available data on unexpected adverse reactions following administration of an investigational medicinal product (IMP) is at odds with the Expert Scientific Group recommendations to expedite the collection, and improve the sharing, of safety information.⁴⁶ Once more, the call by the scientific community for public release of the study results was heard in the days after the BIA 10-2474 trial,^{49,51} yet remains to be fully answered.

Unanticipated immunostimulation caused by parenterally administered pharmaceuticals is not limited to the experimental setting, although there are only a few reports on recalls of drug products as a result of microbial contamination.^{52,53} However, AI may be under-reported because it can be misdiagnosed as hypersensitivity, as occurred in a phase II trial with X.³⁰ Publication bias and other non-scientific reasons may also play a role.

Although these published examples of AI differ from cases X and Y in the type of product and the fact that the cause was contamination instead of an impurity, what they have in common with each other is the fact that that conventional quality control did not generate a safety signal. For example, peptidoglycan, the culprit in cases of aseptic peritonitis,⁵² is not assayed in the standard approach (Table 2).

Alternative assays have been implemented to predict AI with certain pharmaceuticals. Cytokine release assays utilizing human immune cells and other cell types are most commonly used, which have demonstrated proinflammatory responses upon stimulation with several monoclonal antibodies, including TGN1412, rituximab and alemtuzumab.⁵⁴⁻⁵⁸ Host cell impurities or microbial contamination have also been detected with cytokine release assays.^{37,52} The same was true for product X, which elicited an IL-6 and TNF- α response that was completely absent after the production process was modified.²⁹ By contrast, product Y induced no IL-6 or TNF- α release in any of the assays. As many variables can influence the read-out in these assays, and controversy still surrounds the interpretation of the results, cytokine release assays are currently only used as a tool for hazard identification of drug candidates that could induce AI.⁵⁸

Another strategy that has proven useful in detecting immunostimulation employs cell lines transfected with reporter genes for specific components of the innate immune system. Existing cell reporter assays include libraries of human cell lines that are transfected with pathogen pattern recognition receptors (eg, TLR and nucleotide oligomerization domain-like receptors) that are also engineered to express a reporter enzyme (eg, luciferase, or alkaline phosphatase) in response to signal transduction.^{31,59-61} These assays revealed TLR5-mediated activation with product X and TLR4-mediated activation with product Y. Nevertheless, potential synergistical effects of

multiple host cell impurities cannot be detected with this test system. The use of reporter cell lines that endogenously express multiple pattern recognition receptors (such as THP 1 cells) might prove to be useful for this aspect.

Despite the promising developments, no definitive answer to the best preclinical testing strategy can be given, as a fail-safe one probably does not exist. Every laboratory test has its limitations, can suffer from interference and can produce false-negative results, and with these a false sense of safety. Furthermore, the used expression platform, as well as physicochemical and pharmacological properties of the drug substance, each carries different risks of certain impurities being copurified, preventing the use of a general testing approach. However, these limitations should not be used as an argument for maintaining the status quo. Undisputedly, the newer assays described above can detect far more causes of immunostimulation than the currently advised ones (Table 2). Given the potential outcome of missing AI during early drug development, specific testing for AI should be part of performed toxicity studies for biopharmaceuticals, which is in sharp contrast with current guidelines.

It may be considered reassuring that, in retrospect, safety pharmacology studies in cynomolgus monkeys could have detected the immunostimulating propensity of X, provided that the right biomarkers were included. Likewise, signals of AI were observed early in the clinical trials with X and Y. Therefore, investigators should be on the alert for AI when studying biopharmaceuticals in animals and humans, especially if preclinical studies indicated that an IMP may elicit an inflammatory response. As always, a sentinel approach and a cautious dose escalation scheme with a prudently low starting dose should be applied in first-in-human trials. These measures will decrease the likelihood of (severe) AI in later stages, although such reactions can probably never be completely prevented.

Moreover, AIs should not be treated simply as rare idiosyncratic reactions that may be caused by a few biopharmaceuticals. They can severely disrupt vital organ systems, and in many cases indicative signs of AI were left for the discerning eye at some point during development. Awareness is thus probably the most important lesson to be learnt from case X and case Y. Equally important is the publication of cases of AI, to facilitate scientific discussion and improve drug safety.

COMPETING INTERESTS

There are no competing interests to declare.

ORCID

Joannes A. A. Reijers  <https://orcid.org/0000-0002-9532-2357>

REFERENCES

1. Suntharalingam G, Perry MR, Ward S, et al. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *N Engl J Med*. 2006;355(10):1018-1028.
2. US Food and Drug Administration. Immunotoxicology evaluation of investigational new drugs. US Department of Health and Human Services. 2002.

3. Pichler WJ. Adverse side-effects to biological agents. *Allergy*. 2006;61(8):912-920.
4. Palm NW, Rosenstein RK, Medzhitov R. Allergic host defences. *Nature*. 2012;484(7395):465-472.
5. Lieberman P, Garvey LH. Mast cells and anaphylaxis. *Curr Allergy Asthma Rep*. 2016;16(3):20.
6. Chousterman BG, Swirski FK, Weber GF. Cytokine storm and sepsis disease pathogenesis. *Semin Immunopathol*. 2017;39(5):517-528.
7. Behrens EM, Koretzky GA. Cytokine storm syndrome: Looking toward the precision medicine era. *Arthritis Rheumatol*. 2017;69(6):1135-1143.
8. Chung CH. Managing premedications and the risk for reactions to infusional monoclonal antibody therapy. *Oncologist*. 2008;13(6):725-732.
9. Descotes J. Immunotoxicity of monoclonal antibodies. *MAbs*. 2009;1(2):104-111.
10. Scherer K, Spoerl D, Bircher AJ. Adverse drug reactions to biologics. *J Dtsch Dermatol Ges*. 2010;8:411-426.
11. Wilde MI, Goa KL. Muromonab CD3: a reappraisal of its pharmacology and use as prophylaxis of solid organ transplant rejection. *Drugs*. 1996;51(5):865-894.
12. Borchers AT, Leibushor N, Cheema GS, Naguwa SM, Gershwin ME. Immune-mediated adverse effects of biologicals used in the treatment of rheumatic diseases. *J Autoimmun*. 2011;37(4):273-288.
13. Maggi E, Vultaggio A, Matucci A. Acute infusion reactions induced by monoclonal antibody therapy. *Expert Rev Clin Immunol*. 2011;7(1):55-63.
14. Boyman O, Comte D, Spertini F. Adverse reactions to biologic agents and their medical management. *Nat Rev Rheumatol*. 2014;10(10):612-627.
15. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Q6B—Specifications: test procedures and acceptance criteria for biotechnological/biological products. International Council for Harmonisation (ICH). 1999. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q6B/Step4/Q6B_Guideline.pdf
16. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. S6(R1)—Preclinical safety evaluation of biotechnology-derived pharmaceuticals. International Council for Harmonisation (ICH). 2011. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S6_R1/Step4/S6_R1_Guideline.pdf
17. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. S8—Immunotoxicity studies for human pharmaceuticals. International Council for Harmonisation (ICH). 2005. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S8/Step4/S8_Guideline.pdf
18. Gunn GR, Sealey DC, Jamali F, Meibohm B, Ghosh S, Shankar G. From the bench to clinical practice: Understanding the challenges and uncertainties in immunogenicity testing for biopharmaceuticals. *Clin Exp Immunol*. 2016;184(2):137-146.
19. Muller PY, Brennan FR. Safety assessment and dose selection for first-in-human clinical trials with immunomodulatory monoclonal antibodies. *Clin Pharmacol Ther*. 2009;85(3):247-258.
20. Dresser R. First-in-human trial participants: Not a vulnerable population, but vulnerable nonetheless. *J Law Med Ethics*. 2009;37(1):38-50.
21. Committee for Medicinal Products for Human Use (CHMP). Production and quality control of medicinal products derived by recombinant DNA technology. European Medicines Agency. 1994.
22. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Q5A(R1)—Viral safety evaluation of biotechnology products derived from cell lines of human or animal origin. International Council for Harmonisation (ICH). 1999. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q5A_R1/Step4/Q5A_R1_Guideline.pdf
23. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Q5D—Derivation and characterisation of cell substrates used for production of biotechnological/biological products. International Council for Harmonisation (ICH). 1997. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q5D/Step4/Q5D_Guideline.pdf
24. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. S7A—Safety pharmacology studies for human pharmaceuticals. International Council for Harmonisation (ICH). 2000. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S7A/Step4/S7A_Guideline.pdf
25. European Pharmacopoeia Commission. European Pharmacopoeia 2014 (8th edn). Council of Europe: European Directorate for the Quality of Medicines and Healthcare. 2013.
26. United States Pharmacopeial Convention. 2016 USP 39 NF 34 US Pharmacopeia National Formulary, Volumes 1–4. United States Pharmacopeial convention. 2015.
27. WHO Expert Committee on Biological Standardization. Sixty-first report (WHO Technical Report Series). World Health Organization. 2013.
28. Nicholls SJ, Uno K, Kataoka Y, Nissen SE. ETC-216 for coronary artery disease. *Expert Opin Biol Ther*. 2011;11(3):387-394.
29. Reijers JAA, Kallend DG, Malone KE, et al. MDCO-216 does not induce adverse immunostimulation, in contrast to its predecessor ETC-216. *Cardiovasc Drugs Ther*. 2017;31(4):381-389.
30. Nissen SE, Tsunoda T, Tuzcu EM, et al. Effect of recombinant apoA-I Milano on coronary atherosclerosis in patients with acute coronary syndromes: A randomized controlled trial. *JAMA*. 2003;290(17):2292-2300.
31. Huang L-Y, DuMontelle JL, Zolodz M, Deora A, Mozier NM, Golding B. Use of toll-like receptor assays to detect and identify microbial contaminants in biological products. *J Clin Microbiol*. 2009;47(11):3427-3434.
32. Caparon MH, Rust KJ, Hunter AK, et al. Integrated solution to purification challenges in the manufacture of a soluble recombinant protein in *E. coli*. *Biotechnol Bioeng*. 2010;105(2):239-249.
33. Kallend DG, Reijers JAA, Bellibas SE, et al. A single infusion of MDCO-216 (apoA-1 Milano/POPC) increases ABCA1-mediated cholesterol efflux and pre-beta 1 HDL in healthy volunteers and patients with stable coronary artery disease. *Eur Heart J Cardiovasc Pharmacother*. 2016;2(1):23-29.
34. Nicholls S. (2016). Impact of infusion of an apoA-I Milano HDL mimetic on regression of coronary atherosclerosis in acute coronary syndrome patients: The MILANO-PILOT study. *Am. Heart Assoc Sci Sess*. Retrieved from https://professional.heart.org/idc/groups/ahamah-public/@wcm/@sop/@scon/documents/downloadable/ucm_489908.pdf
35. Laude-Sharp M, Haeflner-Cavaillon N, Caroff M, Lantreibecq F, Pusineri C, Kazatchkine MD. Dissociation between the interleukin 1-inducing capacity and limulus reactivity of lipopolysaccharides from gram-negative bacteria. *Cytokine*. 1990;2(4):253-258.
36. Wang X, Morgan DM, Wang G, Mozier NM. Residual DNA analysis in biologics development: Review of measurement and quantitation technologies and future directions. *Biotechnol Bioeng*. 2012;109(2):307-317.

37. Wang X, Hunter AK, Mozier NM. Host cell proteins in biologics development: identification, quantitation and risk assessment. *Biotechnol Bioeng*. 2009;103(3):446-458.
38. Hogwood CEM, Bracewell DG, Smales CM. Measurement and control of host cell proteins (HCPs) in CHO cell bioprocesses. *Curr Opin Biotechnol*. 2014;30:153-160.
39. Dehus O, Hartung T, Hermann C. Endotoxin evaluation of eleven lipopolysaccharides by whole blood assay does not always correlate with limulus amoebocyte lysate assay. *J Endotoxin Res*. 2006;12(3):171-180.
40. Brandenburg K, Howe J, Gutsman T, Garidel P. The expression of endotoxic activity in the limulus test as compared to cytokine production in immune cells. *Curr Med Chem*. 2009;16(21):2653-2660.
41. Copeland S, Warren HS, Lowry SF, Calvano SE, Remick D. Inflammation and the host response to injury investigators. Acute inflammatory response to endotoxin in mice and humans. *Clin Diagn Lab Immunol*. 2005;12(1):60-67.
42. Mestas J, Hughes CCW. Of mice and not men: Differences between mouse and human immunology. *J Immunol*. 2004;172(5):2731-2738.
43. Seok J, Warren HS, Cuenca AG, et al. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci U S A*. 2013;110(9):3507-3512.
44. Brennan FR, Shaw L, Wing MG, Robinson C. Preclinical safety testing of biotechnology-derived pharmaceuticals: Understanding the issues and addressing the challenges. *Mol Biotechnol*. 2004;27(1):59-74.
45. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. S9—Nonclinical evaluation for anticancer pharmaceuticals. International Council for Harmonisation (ICH). 2009. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S9/Step4/S9_Step4_Guideline.pdf
46. Dowsing T, Kendall MJ. The Northwick Park tragedy —Protecting healthy volunteers in future first-in-man trials. *J Clin Pharm Ther*. 2007;32(3):203-207.
47. St Clair EW. The calm after the cytokine storm: Lessons from the TGN1412 trial. *J Clin Invest*. 2008;118(4):1344-1347.
48. Eddleston M, Cohen AF, Webb DJ. Implications of the BIA-102474-101 study for review of first-into-human clinical trials. *Br J Clin Pharmacol*. 2016;81(4):582-586.
49. Mullard A. EMA rewrites phase I guidelines in aftermath of FAAH tragedy. *Nat Rev Drug Discov*. 2016;15:595-595.
50. Kenter M, Cohen A. Establishing risk of human experimentation with drugs: Lessons from TGN1412. *Lancet*. 2006;368(9544):1387-1391.
51. Brøsen K, Funck-Brentano C, Kroemer HK, Pirmohamed M, Schwab M. Open letter on access to the BIA 10-2474 clinical trial data. *Lancet*. 2017;389(10065):156.
52. Martis L, Patel M, Giertych J, et al. Aseptic peritonitis due to peptidoglycan contamination of pharmacopoeia standard dialysis solution. *Lancet*. 2005;365(9459):588-594.
53. Smith RM, Tipple M, Chaudry MN, Schaefer MK, Park BJ. Relapse of fungal meningitis associated with contaminated methylprednisolone. *N Engl J Med*. 2013;368(26):2535-2536.
54. Stebbings R, Eastwood D, Poole S, Thorpe R. After TGN1412: recent developments in cytokine release assays. *J Immunotoxicol*. 2013;10(1):75-82.
55. Moreau T, Coles A, Wing M, et al. Transient increase in symptoms associated with cytokine release in patients with multiple sclerosis. *Brain*. 1996;119(1):225-237.
56. Kamburova EG, van den Hoogen MWF, Koenen HJPM, Baas MC, Hilbrands LB, Joosten I. Cytokine release after treatment with rituximab in renal transplant recipients. *Transplantation*. 2015;99(9):1907-1911.
57. Findlay L, Eastwood D, Ball C, et al. Comparison of novel methods for predicting the risk of pro-inflammatory clinical infusion reactions during monoclonal antibody therapy. *J Immunol Methods*. 2011;371(1-2):134-142.
58. Grimaldi C, Finco D, Fort MM, et al. Cytokine release: A workshop proceedings on the state-of-the-science, current challenges and future directions. *Cytokine*. 2016;85:101-108.
59. Bsibsi M, Bajramovic JJ, Vogt MHJ, et al. The microtubule regulator stathmin is an endogenous protein agonist for TLR3. *J Immunol*. 2010;184(12):6929-6937.
60. Jagessar SA, Kap YS, Heijmans N, et al. Induction of progressive demyelinating autoimmune encephalomyelitis in common marmoset monkeys using MOG34-56 peptide in incomplete Freund adjuvant. *J Neuropathol Exp Neurol*. 2010;69(4):372-385.
61. Haanstra KG, Jagessar SA, Bauchet AL, et al. Induction of experimental autoimmune encephalomyelitis with recombinant human myelin oligodendrocyte glycoprotein in incomplete Freund's adjuvant in three non-human primate species. *J Neuroimmune Pharmacol*. 2013;8(5):1251-1264.

How to cite this article: Reijers JAA, Malone KE, Bajramovic JJ, Verbeek R, Burggraaf J, Moerland M. Adverse immunostimulation caused by impurities: The dark side of biopharmaceuticals. *Br J Clin Pharmacol*. 2019;85:1418-1426. <https://doi.org/10.1111/bcp.13938>