AAPS PharmSciTech (2019) 20:189
DOI: 10.1208/s12249-019-1405-0

Check for updates

Research Article

Nondestructive Quantitative Inspection of Drug Products Using Benchtop NMR Relaxometry—the Case of NovoMix® 30

Katharine T. Briggs,[1] Marc B. Taraban,[1] Weizhen Wang,[2,3] and Y. Bruce Yu[1,4]

Received 11 December 2018; accepted 26 April 2019

Abstract. Batch-level inference-based quality control is the standard practice for drug products. However, rare drug product defects may be missed by batch-level statistical sampling, where a subset of vials in a batch is tested quantitatively but destructively. In 2013, a suspension insulin product, NovoLog® Mix 70/30 was recalled due to a manufacturing error, which resulted in insulin strength deviations up to 50% from the labeled value. This study analyzed currently marketed FlexPen® devices by the water proton transverse relaxation rate using a benchtop nuclear magnetic resonance relaxometer. The water proton transverse relaxation rate was found to be sensitive to detecting concentration changes of the FlexPen® product. These findings support the development of vial-level verification-based quality control for drug products where every vial in a batch is inspected quantitatively but nondestructively.

KEY WORDS: NMR; insulin; quality control; nondestructive inspection; adverse events.

INTRODUCTION

On September 27, 2013, a woman was found collapsed on the floor of her home. According to the FDA MAUDE Adverse Event Report database (catalog number 185230) (1), she suffered from excessive thirst, a collapsed middle lobe of her right lung, went into a coma, was treated at a hospital, recovered, and later was discharged. Prior to the incident, the patient self-administered NovoMix® 30 Penfill®. Tests indicate that her insulin dose was incorrect. Misdosing of insulin for diabetic patients could cause hypoglycemia or hyperglycemia, with consequences as serious as coma, or even death. Hypoglycemia caused by misdosing of diabetic drugs, including insulin, is rather common (2,3).

Possible causes for insulin misdosing in the event on Sept. 27, 2013 could have been user error (*i.e.*, the woman dialed the wrong dose), mechanical malfunction (*i.e.*, the Penfill® device delivered the wrong dose), or manufacturing error (*i.e.*, Penfill® contained the wrong dose). Four weeks after the incident, Novo Nordisk, the company that makes NovoMix® 30, issued a recall of 30 batches of the NovoMix® 30 FlexPen® and 3 batches of the NovoMix® 30 Penfill® (4). In total, 3.3 million units of FlexPen® and Penfill® devices were recalled (5). In this paper, we will use "vial," instead of "unit," as a generic term for any drug product presentation, including vials, bottles, bags, syringes, and pens. This is primarily to avoid confusion with insulin dose level; according to the product label, each FlexPen® contains 100 units/mL of insulin.

The reason for the recall is that an error in cartridge filling caused the insulin dose in some cartridges to deviate by up to ± 50% of the specified value, 10 times larger than the ± 5% potency deviation permitted by the US Pharmacopeia (6). The error rate was 0.14% (4). For 3.3 million recalled vials, this translates into 4620 defective vials. The total number of defective vials that have been used by patients is unknown

Key Points

- Low-field benchtop NMR relaxometry is capable to detect drug dose deviations in intact drug products by measuring the water proton transverse relaxation rate.
- Measuring the water proton relaxation rate of a drug product is a nondestructive method that has the potential to serve as a quality control method for released drug products.
- Prime candidates for nondestructive verification-based quality control are scenarios where the occurrence of drug product defects is low, but the consequences of a defect are serious.

Electronic supplementary material The online version of this article (https://doi.org/10.1208/s12249-019-1405-0) contains supplementary material, which is available to authorized users.

[1] Department of Pharmaceutical Sciences and Bio- and Nano-Technology Center, University of Maryland, 20 Penn Street, Baltimore, Maryland 21201, USA.
[2] Department of Statistics, Tianjin University of Finance and Economics, Tianjin, 300222, People's Republic of China.
[3] Department of Mathematics and Statistics, Wright State University, Dayton, Ohio 45435, USA.
[4] To whom correspondence should be addressed. (e–mail: byu@rx.umaryland.edu)

Abbreviations: *NMR*, Nuclear magnetic resonance; T_2, Transverse relaxation time; R_2, Transverse relaxation rate; *QC*, Quality control; *DP*, Drug product; *DS*, Drug substance; *NS*, Number of scans; *RD*, Relaxation delay; *S/N*, Signal/noise.

Published online: 20 May 2019

and probably unknowable. However, it is known that the Penfill® used by the woman belong to one of the recalled batches of NovoMix® 30 Penfill® (batch no. cs6d422). Therefore, the likely cause of this adverse event was the wrong insulin dose in the product. It is impossible to know this with certainty because the product used by the woman prior to the adverse event was not returned for examination, according to the FDA MAUDE database.

In the US market, biologics, such as insulin, are associated with more safety events (7) and have higher recall rates (8) than small-molecule drugs. The woman's collapse on Sep. 27, 2013, was only one of many adverse drug events each year (1), and the recall of NovoMix® 30, 1 month later, was only one of many drug recalls each year (9). This episode is typical in another way; there was no data on the specific Penfill® involved in the adverse event, making it impossible to draw a definitive causal link between an adverse event and a product. The FDA MAUDE database states "Novo Nordisk evaluates the causality between the events due to low blood sugar and the suspected products as possible." Possible, not definite.

As pointed out in a recent paper by FDA scientists, pharmaceutical manufacturing is far from six-sigma quality (10), which means at most 3.4 errors per one million opportunities. Considering the manufacturing of biologics has many steps, each with multiple opportunities for errors, it is inevitable some vials of the final product will be defective. It is up to product inspection to catch and remove defective vials at the finish line. In fact, waste in the form of discarded products could reach 50% in some cases (11). Existing product inspection apparently failed in the NovoMix® 30 case.

Benchtop time-domain NMR techniques have been used for drug product quality control, such as content uniformity of tablets containing fluorinated active pharmaceutical ingredients through the ^{19}F signal intensity (12) and filling uniformity of solids and liquids in vials through the proton (^{1}H) signal intensity (13). In this work, we use the $^{1}H_2O$ transverse relaxation rate $R_2(^{1}H_2O)$ to access content uniformity of insulin pens and to explore its potential to quantitatively, yet nondestructively, inspect every vial in a batch of a drug product. For NovoMix® 30, a 10% deviation from the insulin level specified in the drug label was reliably detected using a benchtop NMR device at 0.5 T, and the inspection time for each vial was 30–60 s.

METHODS

Instrument for Nondestructive Inspection

A benchtop NMR Analyzer (MQC+ from Oxford Instruments) was used for drug product inspection. The magnetic field strength was 0.5 T, with a ^{1}H resonance frequency of 23.8 MHz. The magnet temperature was 25°C. The 1193 Pro 26 mm probe readily accommodates vials with an outer diameter of 25 mm or less.

Drug Product

The NovoLog® Mix 70/30 FlexPen®, purchased in the US market to match the European NovoMix® 30, was used in this study, which is the same product recalled after the above incident (1). A carton of five NovoLog® Mix 70/30 FlexPen® devices (Cardinal Health) was procured; the cartridge of each FlexPen® contains 3 mL of 70% protamine-crystallized insulin aspart and 30% soluble insulin aspart (100 units/mL or 3.5 mg/mL). The cartridge containing the whitish insulin suspension is encased inside a blue nontransparent plastic shell, forming the pen body with an outer diameter of 15.5 mm (Fig. S1).

Recreating the Defect in 2013 NovoMix® 30 Recall

Pens were labeled 1 through 5 and FlexPen®#3 was used for these dose deviation experiments. Recreation of negative insulin level deviation (from 100 down to 50% of the original insulin suspension) involved drawing a portion of the insulin aspart suspension out of the FlexPen® and replacing it with the same volume of FlexPen® buffer, while both the pen and buffer were kept on ice to maintain 2–8°C (Table I).

The FlexPen® buffer was prepared according to the inactive ingredients listed in the product label (package insert) for NovoLog® Mix 70/30 (14): glycerol 16.0 mg/mL, phenol 1.5 mg/mL, meta-cresol 1.72 mg/mL, zinc 19.6 μg/mL, disodium hydrogen phosphate dihydrate 1.25 mg/mL, sodium chloride 0.877 mg/mL, and protamine sulfate 0.320 mg/mL. The buffer was 0.2 μm filtered prior to adding the zinc dust and was pH adjusted to 7.2–7.4 using hydrochloric acid or sodium hydroxide.

The FlexPen® drug substance was gradually diluted to achieve a lower insulin concentration, 50% of the original suspension, by ~ 10% dilution at each step. We were unable to recreate positive insulin level deviations (from 100 up to 150% of the original insulin suspension) because we do not have access to insulin aspart powder. The dilution of FlexPen®#2 content was transferred to and measured in glass vials (21 mm × 50 mm, 3 dram vials, Kimble Chase) following the same procedure for dilution as in FlexPen®#3.

Before drawing the insulin suspension out, the FlexPen®#3 was rolled between the hands horizontally and inverted vertically, 10 times each, to ensure that the insulin particles were fully suspended according to the product label. Then, a portion of the insulin suspension was removed using a 3-mL syringe and 23 gauge needle and then dispelled into a microcentrifuge tube. The exact volume removed was confirmed by pipettor. This volume of insulin suspension was replaced by equal volume of buffer, which was injected to the FlexPen® using a syringe and needle. Next, the FlexPen® was inspected by NMR (see "Nondestructive Inspection of Defective Drug Product") and the dilution process was repeated again to gradually achieve a lower relative insulin concentration of 4.1%. During this dilution procedure, the FlexPen® was kept on ice, and the buffer was chilled on ice. The following is a table of the dilution procedure, and the actual pipetted amounts (Table I).

Nondestructive Inspection of a Drug Product

Nondestructive quantitative inspection refers to measuring the water proton ($^{1}H_2O$) transverse relaxation time ($T_2(^{1}H_2O)$) of each FlexPen®, which contains an aqueous suspension of soluble and insoluble insulin aspart particles.

Table I. Serial Dilution of Insulin Suspension in FlexPen® #3.

Sample	Concentration before each serial dilution step (%)[a]	Volume replacement (mL)[b]	Concentration after dilution (%)[c]
Dilution 1	100	0.302	89.93
Dilution 2	89.93	0.333	79.95
Dilution 3	79.95	0.373	70.01
Dilution 4	70.01	0.431	59.95
Dilution 5	59.95	0.500	49.96
Dilution 6	49.96	1.422	26.28
Dilution 7	26.28	1.422	13.82
Dilution 8	13.82	1.475	7.03
Dilution 9	7.03	1.250	4.10

[a] 100% concentration is equal to 3.5 mg/mL
[b] Volume replacement refers to the volume of insulin suspension in a FlexPen® removed and replaced by buffer. The total volume of insulin suspension in the FlexPen® was kept at 3 mL throughout the dilution process
[c] Concentration% after dilution = (Concentration% before dilution) × (3 − volume replacement)/3

$T_2(^1H_2O)$ was measured using a standard time-domain CPMG (Carr-Purcell-Meiboom-Gill) (15) pulse sequence of RD-90°-τ-180°-τ, where RD was a 12 s relaxation delay, 90° and 180° are rf pulses, and τ is the inter-pulse delay of 500 μs, and 9000 echoes were collected. Each measurement began with one dummy scan to ensure that the subsequent two scans (number of scans (NS)) of acquired data are of a spin system that has reached a steady state. The 90° and 180° pulses were calibrated initially and then used in all of the measurements. The offset (O1) and detection phase (DP) were calibrated just prior to each $T_2(^1H_2O)$ measurement. To determine the $T_2(^1H_2O)$ and the standard error, the intensities of the 9000 echoes were plotted in WinFit (Oxford Instruments) and fit to a single exponential decay (Eq. 1 in text). The average T_2 of pure water, measured with three different flip angles (90°, 45°, and 20°), was 2.567 ± 0.006 s, which suggests no evidence of radiation damping (otherwise, the 3 flip angles would have very different T_2 values). We, therefore, used a 90° flip angle in these experiments because we do not suspect radiation damping is affecting the measurement.

$T_2(^1H_2O)$ was measured for five FlexPen® devices consecutively, one at a time, over a stretch of ~ 20 min for all. To assess experimental error, $T_2(^1H_2O)$ was measured five times for each FlexPen® over a 5-week period, i.e., one measurement per week for each pen. Between inspections, FlexPen® devices were stored at 2–8°C, in accordance with the storage condition specified in the product label. During the 20 min span of measurements, the FlexPen® devices were kept in a Styrofoam insulation box with a refrigeration pack atop the FlexPen® carton and a frozen pack atop the refrigeration pack, for an approximate incubation at 2–8°C, until the measurement time. In an 18°C temperature-controlled NMR room, at the measurement time, a FlexPen® was rolled and inverted, according to the manufacturer's instructions for suspension of the insulin aspart. For ease of lowering into and lifting out of the bore, the FlexPen® was put inside a flat bottom glass tube (26 mm o.d.,22.8 i.d., 200 mm length, Oxford Instruments), which was then inserted into the magnet bore without its cap removed (Fig. 1). The CPMG experiment was as described above. For all FlexPen® devices, the time from the Styrofoam insulation box and carton to the start of the experiment was approximately 1 min and 30 s, and the time of the actual inspection was ~ 1 min.

$T_2(^1H_2O)$ was extracted from the time decay of the echo intensity (I(t)) of the ^1H_2O signal according to the following equation for single exponential decay:

$$I(t) = I_o \times e^{-\frac{t}{T_2}} \quad (1)$$

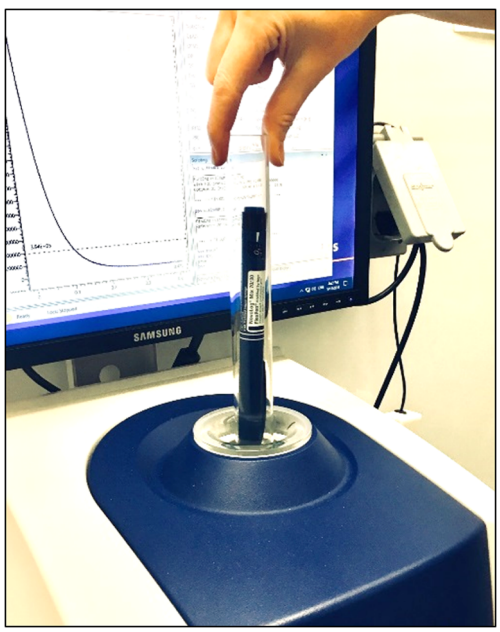

Fig. 1. Inspection of intact FlexPen®. The pen was being lowered into the bore of the benchtop NMR Analyzer for inspection. The drug substance, a whitish aqueous suspension of insulin, is contained in a cartridge inside the FlexPen®

Nondestructive Inspection of Defective Drug Product

The $T_2(^1H_2O)$ of intentionally created defective FlexPen® devices were also measured in the NMR Analyzer, as shown in Fig. 1 using the same procedure and NMR parameters as the original FlexPen® product (see "Nondestructive Inspection of a Drug Product") .

Experiments for Faster Inspection

Inspections, using the CPMG pulse sequence (described in "Nondestructive Inspection of a Drug Product" and "Nondestructive Inspection of Defective Drug Product"), take ~ 60 s. The T_1 of the FlexPen® buffer is 1.622 s. The goal of these experiments was to push the boundaries to determine the intersection of minimal experiment time and high quality. To explore whether reduction of the inspection time was feasible, two inspection parameters, NS and RD, were reduced respectively from 2 to 1 and from 12 to 6 s. Three sets of parameters with (NS, RD) = (2, 12 s), (2, 6 s), and (1, 6 s) were tested. The (NS, RD) = (2, 12 s) set was used in the original inspections ("Nondestructive Inspection of a Drug Product" and "Nondestructive Inspection of Defective Drug Product") and was repeated here, 5 months after the original inspection, to verify the stability of the product and the reproducibility of the inspection.

The intact undiluted FlexPen®#4 was used for these experiments. It was inspected thrice daily, corresponding to three separate parameter sets, for six nonconsecutive days within the span of 1.5 weeks. FlexPen®#4 was kept in a Styrofoam insulation box on ice in a plastic bag directly before and after the measurement. As before ("Nondestructive Inspection of a Drug Product" and "Nondestructive Inspection of Defective Drug Product"), prior to the measurement, the FlexPen®#4 was rolled and inverted between the hands for 15 s in an 18°C room, where the benchtop NMR is located. The FlexPen® was placed inside the magnet bore (25°C) where the parameters were set and loaded (~ 45 s). Subsequently, the $T_2(^1H_2O)$ data were acquired and measurement times varied between 30 and 60 s, depending on the measurement parameters (NS and RD).

Statistical Sampling Calculation

We calculated the probability of rejecting a batch of a product with a defect rate of 0.14% through acceptance sampling. In acceptance sampling, there are two parameters related to inspection outcome, the acceptable quality limit (AQL) and the rejectable quality limit (RQL), with $0 \le \text{AQL} < \text{RQL} \le 1$. AQL and RQL are set by the manufacturer and depend on product/defect types. Let p be the true defective rate of drug vials in a batch of size N. The decision rule is:

$$\text{accept when } p \le \text{AQL}; \text{reject when } p \ge \text{RQL} (\text{Rule 1}).$$

However, when destructive testing is performed, the true defective rate p is unobservable because only a fraction of vials in a batch is tested. Therefore, Rule 1 cannot be executed.

Suppose n vials are selected from a batch of N vials for destructive testing ($1 \le n \le N$), and it is found that X of the n vials are defective ($0 \le X \le n$). Then, the apparent defect rate $\hat{p} = X/n$. Note that the apparent defect rate $\hat{p}$ may or may not equal the true defect rate p, which remains unknown after the testing (unless $n = N$). The new decision rule, which is implemented in terms of X rather than $\hat{p}$, is:

$$\text{accept when } X < x_0; \text{reject when } X \ge x_0 \text{ (Rule 2)}$$

where $\{X \ge x_0\}$ is called the rejection region for the hypothesis: $p \le \text{AQL}$, and x_0 is calculated in Eq. (3) below.

Rule 2 based on X and x_0, instead of Rule 1 based on p, AQL, and RQL, may commit two types of error: the producer's risk α and the buyer's risk β. The producer's risk α is the probability that a conforming batch (a good batch) is wrongly rejected, $\alpha = P(X \ge x_0 | p = \text{AQL})$; the buyers risk β is the probability that a nonconforming batch (a bad batch) is wrongly accepted, $\beta = P(X < x_0 | p = \text{RQL})$. The root cause of the errors is because the true effect rate p remains unknown after limited destructive testing. $1 - \beta$, called power, is the probability that a nonconforming batch (a bad batch) is rightly rejected.

It can be shown X follows a hypergeometric distribution, denoted by Hyper(pN, N, n), *i.e.*, its probability mass function is:

$$P(X = x) = \frac{\binom{pN}{n}\binom{N-pN}{n-x}}{\binom{N}{n}} \tag{2}$$

for any integer x between $\max\{0, n + pN - N\}$ and $\min\{n, pN\}$.

In typical acceptance sampling, α and β are respectively set to 0.05 and 0.10, and then n and x_0 are determined so that Rule 2 can be executed. In this work, however, our intent is to demonstrate how the probability to rightly reject a bad batch, $1 - \beta$, varies with the sample size n. To this end, we set α at 0.05, vary n from 1 to N, and calculate $1 - \beta$ for each n.

From Eq. 2 and the producer's risk α, x_0 is given by:

$$x_0 = min\left\{ x : \alpha = P(X \ge x | p = AQL) = \sum_{y=x}^{n} \frac{\binom{AQL*N}{y}\binom{N-AQL*N}{n-y}}{\binom{N}{n}} \le 0.05 \right\}. \tag{3}$$

Then, β is calculated from the following equation:

$$\beta = P(X < x_0 | p = RQL) = \sum_{y=0}^{x_0} \frac{\binom{RQL*N}{y}\binom{N-RQL*N}{n-y}}{\binom{N}{n}}. \tag{4}$$

An R-code is available from one of us (W.W.) upon request to compute x_0 and β when N, n, AQL, RQL, and the nominal level of α are given.

The 2013 recall of NovoMix® 30 involves 3.3 million vials spread over 33 batches (5), which gives an average batch size of 100,000. Hence, we set $N = 100{,}000$. The recalled batches have a defect rate of 0.14% (note: it is unclear whether 0.14% is the true defect rate p or the apparent defect rate $\hat{p}$); hence, we set RQL = 0.0014. The recall did not specify AQL. For our calculation, we set AQL = 0.001. Note that setting AQL = 0 would mean that if a single defective vial is found in a batch during acceptance sampling, the entire batch would have to be rejected.

RESULTS

In contrast to high-field NMR spectroscopy where the proton signals can be separated by frequency differences, benchtop NMR relaxometry in the time domain measures all the protons in one single signal. Benchtop NMR was explored as a tool for nondestructive and quantitative inspection of a drug product *via* the water proton (^1H_2O) NMR signal. In aqueous liquid dosage forms (solutions, suspensions, and emulsions), water is by far the dominate component. The drug product that we used was a FlexPen®, referred to, here, as an insulin pen or simply a pen. Based on our prior work (16,17), we chose the water proton transverse relaxation rate $R_2(^1H_2O)$ as the parameter to assess insulin concentration in the insulin pens. If the five pens have the same insulin concentration, they should, within experimental error, have the same $R_2(^1H_2O)$ value.

Variability of $R_2(^1H_2O)$ Measurement

The transverse relaxation rate, $R_2(^1H_2O)$ $(= 1/T_2(^1H_2O))$, of five pens, numbered 1 through 5, was measured once a week for 5 weeks. From these five $R_2(^1H_2O)$ values of each pen, the measurement variability can be determined, which ranges between 0.002 to 0.012 s^{-1} for different pens, with an average variability of 0.008 s^{-1} (Table II). For raw $T_2(^1H_2O)$ data, see Table S1 and Fig. S3. There are several sources for this variability, such as instrument stability, power supply stability, room temperature stability, and sample (insulin pen) temperature. These measurements are meant to be interpreted relative to each other (*i.e.*, as a consistency check) because they have the same rolling/inverting mixing protocol and approximately the same amount of temperature rising of the samples over the course of the measurement but are not measured at thermal equilibrium with the NMR probe.

Assessing Insulin Filling Evenness in FlexPen®

To assess insulin filling evenness in the five pens from the same carton, the $R_2(^1H_2O)$ of each pen was measured back-to-back within ~20 min. This was repeated five times, spreading over 5 weeks. The pen-to-pen variability on the same day is in the range of 0.006 to 0.008 s^{-1}, with an average of 0.007 s^{-1}, comparable to the measurement variability of 0.008 s^{-1} (Fig. S3, Table II, Table S1). From this, we conclude that the five pens were filled evenly with the insulin suspension. However, $R_2(^1H_2O)$ can measure only the relative but not the absolute insulin level, which is specified in the product label as 100 units/mL, equivalent to 3.5 mg/mL (6). In theory, it is possible that all five pens could have the same incorrect insulin concentration (*e.g.*, 2 mg/mL), *i.e.*, they were filled evenly but incorrectly, but this is unlikely.

Detecting Uneven Insulin Filling in Insulin Pens

We explored whether uneven insulin filling in pens can be detected by the $R_2(^1H_2O)$. To this end, we first needed to recreate uneven filling, which was up to ± 50% deviation from the specified value of 3.5 mg/mL in the 2013 NovoMix® 30 recall. That translates into an insulin concentration range of 1.75–5.25 mg/mL. Without access to the insulin aspart powder, we could not make the insulin concentration higher than 3.5 mg/mL. However, we made the insulin concentration lower than 3.5 mg/mL by diluting the original insulin suspension with the same buffer used by the drug maker in the insulin suspension formulation.

$R_2(^1H_2O)$ of FlexPen® #3 filled with insulin suspensions in the concentration range of 1.75–3.5 mg/mL were measured (Table S2). $R_2(^1H_2O)$ increases linearly with the insulin concentration, with a slope of 0.0726 $s^{-1}\cdot(mg/mL)^{-1}$ (Fig. 2). The same slope was obtained when $R_2(^1H_2O)$ of the insulin suspensions from FlexPen® #2 was measured in glass vials instead of pens although the absolute value of $R_2(^1H_2O)$ differs slightly for pens and glass vial (Figs. S2, S4). Hence, the slope of 0.0726 $s^{-1}\cdot(mg/mL)^{-1}$ is not a characteristic of the drug product (insulin suspension plus pen container) but rather a characteristic of the drug substance (insulin suspension).

Although there are no $R_2(^1H_2O)$ data for positive insulin concentration deviations (*i.e.*, > 3.5 mg/mL), the linear growth of $R_2(^1H_2O)$ with solute concentration has been observed in several protein, peptide, and amino acid systems over concentration ranges much wider than the 1.75–3.5 mg/mL range in this work (16,18). Therefore, it is reasonable to assume that the same linear relationship and slope will hold for pens with positive insulin level deviation (3.5–5.25 mg/mL), provided the suspended insulin particles do not aggregate in this concentration range.

With a slope of 0.0726 $s^{-1}\cdot(mg/mL)^{-1}$, 5% dose deviation from 3.5 mg/mL translates into $R_2(^1H_2O)$ change of 0.013 s^{-1}, slightly above the measurement error of 0.009 s^{-1} (Table II). Note US Pharmacopeia permits 5% potency deviation (6) while the dose deviation in the NovoMix® 30 recall was up to 50%. The data collected in this work suggest 10% dose deviation in insulin pens can be reliably detected by a benchtop NMR Analyzer at 0.5 T. With more optimized measurement parameters, sufficient temperature equilibration, and longer data collection time, 5% dose deviation might also be detected reliably.

Reducing the Inspection Time

To explore whether faster inspection can be achieved without sacrificing the quality of the inspection, the NS and RD (in seconds) in the CPMG pulse sequence were reduced. The intact FlexPen® #4 was used for this purpose.

Generally, the greater the NS, the higher the signal to noise (S/N) ratio will be. As for RD, it should be larger than $5T_1$ (T_1 is the longitudinal relaxation time). The criteria are whether the echo intensity of the ^1H_2O signal is high enough

Table II. FlexPen® Inspection $R_2(^1H_2O)$ data and variability (all in s^{-1}). Each $R_2(^1H_2O)$ is calculated from corresponding $T_2(^1H_2O)$ (Table S1) using $R_2(^1H_2O) = 1/T_2(^1H_2O)$.

	Week 1	Week 2	Week 3	Week 4	Week 5	Measurement variability
Pen #1	0.972	0.993	1.000	0.991	1.003	0.012
Pen #2	0.980	0.992	1.003	0.992	1.003	0.009
Pen #3	0.972	0.993	0.996	1.001	1.001	0.012
Pen #4	0.986	1.002	1.000	0.988	0.993	0.007
Pen #5	0.986	0.982	0.986	0.986	0.985	0.002
Pen-to-pen variability	0.007	0.007	0.007	0.006	0.008	*0.009*

(i) Inspection was made once a week for each FlexPen® for 5 weeks; (ii) measurement variability is the standard deviation of the row of single $T_2(^1H_2O)$ measurements made each week on the same pen, over 5 weeks; (iii) Pen-to-pen variability is the standard deviation of the column of single $T_2(^1H_2O)$ measurements made for five pens on the same day; (iv) The combined average (0.009 s^{-1} in italic) is based on 25 $R_2(^1H_2O)$ data points (*i.e.*, from five pens over 5 weeks)

that the reduction in NS does not significantly decrease the S/N ratio and whether a reduced RD is still long enough for the ^1H_2O echo signal intensity to fully vanish to zero prior to the next scan (Fig. S5). Low S/N ratio or incomplete relaxation may result in deceptive or unreliable results and would not be conducive to catching a defective drug product.

As shown in Table S3, reducing NS from 2 to 1 and RD from 12 to 6 s halves the inspection time, from 64 to 32 s, with negligible impact on $T_2(^1H_2O)$ data quality (1.008 ± 0.012 s *vs.* 1.013 ± 0.009 s). Further, $T_2(^1H_2O)$ data collected on FlexPen® #4 5 months apart displays a negligible difference (1.008 ± 0.012 s *vs.* 1.006 ± 0.007 s), attesting to the reproducibility of the inspection as well as the stability of the product, when stored properly (2–8°C).

Probability of Catching Rare Product Defects Using Statistical Sampling

Typical quality control (QC) of parenteral drug products involves visual inspection and destructive testing (acceptance sampling). Visual inspection is qualitative and is performed on all vials in a batch, sometimes more than once (17). Destructive testing is quantitative and is performed on a subset of vials, typically 10–20, in a batch (17). From the quality of the tested vials, a decision is made to accept or reject the untested vials in the same batch.

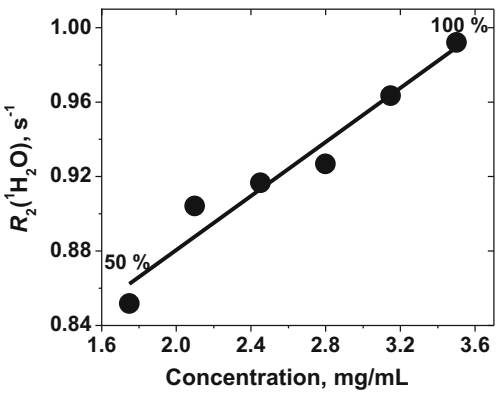

Fig. 2. FlexPen® dilution measured by water relaxation rate. The change in FlexPen® concentration, from 1.75 to 3.5 mg/mL insulin, is readily detected by the $R_2(^1H_2O)$. Each circle represents a 10% change in concentration from 100 to 50%. The slope of the linear fit is 0.0726 ± 0.008 $s^{-1} \cdot (mg/mL)^{-1}$ ($R^2 = 0.943$)

When a drug product defect is not visually obvious (such as wrong insulin concentration caused by uneven filling) and the defect rate is low, it poses a challenge to current QC practice. Table III presents the probability to catch a defect with a rate of 0.14% when *n* vials in a batch of 100,000 vials are subjected to destructive testing. When only 10–20 vials out of 100, 000 vials are tested, as is typical (17), the chance to reject a nonconforming batch with a defect rate of 0.14% is less than 3%, which is problematic from the drug safety standpoint. To reach 90% chance to reject such a nonconforming batch, ~ 40,000 out of 100,000 vials need to be destructively tested, which is prohibitive from the drug cost standpoint. While the exact probability values for statistical sampling to catch or miss product defects may vary from product to product, the conflict between drug safety and drug cost is inherent to destructive testing.

DISCUSSION

NovoMix® 30 was first marketed in the European Union in 2000 and in the USA in 2001. By 2013, the drug maker had over 10 years of experience in making this product at commercial scale. The error that led to the NovoMix® 30 recall occurred at the fill-finish step of the manufacturing process; cartridges were filled unevenly with the drug substance. NovoMix® 30 is a suspension product due to the presence of insoluble insulin. Compared with solutions, suspensions are heterogeneous and have complex hydrodynamic and flow behaviors (19,20), which make them susceptible to uneven filling. In fact, uneven filling may occur not only to suspensions but also to concentrated protein solutions because of their high viscosity (21–23). Uneven filling of drug substance (DS) into vials creates defects at the drug product (DP = DS + sealed and labeled vial) level with vial-to-vial variability.

Challenges to Drug Product Quality Control

Defects at the DS level (*e.g.*, degradation/modification of active/inactive ingredients) are easier to catch because

Table III. The probability to wrongly reject a conforming batch (α) and the probability to rightly reject a nonconforming batch ($1-\beta$) as functions of the number of tested vials n. The batch size N is 100,000. The batch is rejected if at least x_0 units are found nonconforming. The acceptable quality limit (AQL) and rejectable quality limit (RQL) are respectively 0.1% and 0.14%. The small values of α and $1-\beta$ at $n = 400$ are due to the highly skewed nature of the hypergeometric distribution function, from which α and $1-\beta$ are calculated.

n	4	10	20	40	400	4000	40,000	100,000
x_0	1	1	1	1	3	8	49	101
α	0.004	0.010	0.020	0.039	0.008	0.047	0.042	0
$1-\beta$	0.006	0.014	0.028	0.055	0.019	0.199	0.903	1

one has the entire armamentarium of analytical techniques to deploy. At the DP level, one has very limited options to catch defects because the DS is already sealed inside the vial; drawing the DS out of the vial for analysis amounts to tampering and is, thereby, destructive from the standpoint of DP integrity. Typical QC procedures at the DP level involve qualitative visual inspection of every vial in a batch, followed by quantitative destructive testing of a few vials in a batch (acceptance sampling) (17).

Sampling-based QC procedures accept, reject, and recall DPs at the batch level, not the individual vial level because there is no data at the individual vial level (rejection of individual vials may occur at the visual inspection step, which is typically before destructive sampling). In fact, a released vial has no quantitative QC data on it. When it comes to rare but serious DP defects, this practice has at least three downsides. First, rare defects may escape detection and harm patients, as illustrated by the case of NovoMix® 30. Second, when rare defects are caught, many conforming vials are rejected or recalled along with the few nonconforming vials. For the NovoMix® 30 recall, 3,295,380 conforming vials were recalled alongside 4620 nonconforming (defective) vials. This indiscrimination between conforming and nonconforming vials in product rejection/recall is not only wasteful but may also contribute to drug shortages (24). Third, when adverse events occur, it is very difficult to draw a definitive link between the drug product and the adverse event. The reason is simple: there is no data on the vial taken by the patient who developed the adverse reaction. In many cases, the DP (or the DS inside container) taken by the patient is forever gone, making retroactive data collection impossible.

If a recalled batch is analyzed again using statistical sampling, rare defects, which evaded prerelease QC, may evade detection again. If this were to occur, drug maker and/or health authority might conclude that there is no evidence of a causal link between the adverse event and the drug product, then the recalled batch might be reintroduced into the market and might cause harm again. The chance that a specific patient will be harmed by a rare product defect is low, but the chance that some patient somewhere will be harmed by a rare product defect may actually be quite high. After all, adverse events do occur, and some are never explained. Is it possible that some unexplained adverse events are caused by rare product defects? Without quantitative data at the individual vial level, it is impossible to answer this question definitively (25).

Enhanced Drug Product QC Through 100% Nondestructive Quantitative Inspection

One way to overcome the inherent limitation of sampling-based QC practice is to perform nondestructive quantitative inspection on every vial in a batch (*e.g.*, every NovoLog® 70/30 FlexPen®). The function of the nondestructive quantitative inspection at the DP level is to ensure product uniformity (*e.g.*, all pens have the same insulin dose) rather than to ascertain absolute product quality (*e.g.*, all pens contain 3.5 mg/mL insulin). The absolute product quality still needs to be established by destructive quantitative analysis of a few vials in each batch. Having vial-level quantitative data will make accept/reject decisions much more precise and will likely reduce the chance of adverse drug events caused by released defective drug products, thereby averting some drug recalls. In the case of an adverse event, data on the specific vial taken by the patient can help inform whether there is a causal link between the product and the adverse event. In the case of defective vials identified in recalled products, data collected at different time points in the lifetime on the defective vials could help pinpoint where/when/how the defects occurred. Currently, reasons for product failures may remain unknown even after investigation (11).

From the statistical point of view, the destructive sampling-based QC practice uses a subpopulation to infer the entire population, and an incorrect inference may occur due to a limited sample size. In contrast, 100% nondestructive inspection, similar to a census, would investigate the entire population. Its conclusion does not involve making inference about the entire population from data collected on a subpopulation. Obviously, 100% nondestructive inspection would be preferred, whenever feasible.

Enabling Technologies for 100% Nondestructive Quantitative Inspections

One hundred percent nondestructive quantitative inspection requires enabling technologies. In this work, we demonstrated that *w*NMR has this potential. In the case of NovoMix® 30, *w*NMR can detect insulin level deviation in approximately 30 s, comparable to the time it takes to visually inspect a vial in a batch. At this rate, the throughput of *w*NMR is comparable to that of human visual inspection.

One could envision, in the future, that drug products would have a $R_2(^1H_2O)$ value or range, which would be considered acceptable for passing as a safe drug product. With a quick *w*NMR inspection, any major deviations from the specified value or range would be detected at the time of product release, in the distribution chain, and at the point of care. The inspection results could be reported back to relevant parties (drug maker, distributor, health authority, *etc.*) as part of a pharmacovigilance program.

A potential starting point to implement 100% quantitative inspection might be with vaccines that contain aluminum adjuvants (*e.g.*, Prevnar 13®, Gardasil®, and Pediarix®). Aluminum adjuvants are micron-sized particles that are heavier than water and tend to sediment, which

may cause uneven filling and vial-to-vial variation of aluminum content. Overfilling might make the aluminum content higher than the permitted level (0.85 mg/dose, according to the US Code of Federal Regulations (610.15(a))) while underfilling might reduce vaccine effectiveness. One hundred percent quantitative inspection on aluminum-adjuvanted vaccines could help to ensure that the aluminum level in every vial of these vaccines indeed lies within the acceptable range, agreed upon by vaccine manufacturers and regulators, to ensure vaccine safety and efficacy.

CONCLUSION

As shown by the NovoMix® 30 case in 2013, uneven filling of drug substances into vials during manufacturing may cause serious product defects and harm patients. Current QC practice, which rests on destructive sampling of a small number of vials in a batch, may fail to detect filling errors. Many complex drugs and vaccines are suspensions and emulsions with complex flow properties, which increase the likelihood of filling errors. One hundred percent nondestructive quantitative inspection of such products might be prudent. Indeed, an important aspect of drug product QC is to manage uncertainty (26). Lack of vial-level quantitative data on released drug products is a source of uncertainty. While uncertainty may never be fully eliminated, 100% quantitative inspection can reduce uncertainty.

FUNDING SOURCE

K.T.B. received an M-CERSI Postdoctoral Fellowship from the FDA. A part of the work was supported by the National Institute for Innovation in Manufacturing Biopharmaceuticals (NIIMBL). The University of Maryland Baltimore provided additional funding.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest K.T. Briggs, M.B. Taraban, and Y.B. Yu are co-inventors on issued and pending patents on wNMR technologies.

Ethical Approval This article does not contain any studies with human or animal subjects.

REFERENCES

1. U.S. Food and Drug Administration (FDA). MAUDE adverse event report: Novo Nordisk A/S Novopen insulin delivery device 2013. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/detail.cfm?mdrfoi__id=3621493. Accessed 09 Feb 2018.
2. Ben Salem C, Fathallah N, Hmouda H, Bouraoui K. Drug-induced hypoglycaemia. Drug Saf. 2011;34:21–45. https://doi.org/10.2165/11538290-000000000-00000.
3. McCall AL. Insulin therapy and hypoglycemia. Endocrinol Metab Clin N Am. 2012;41:57–87. https://doi.org/10.1016/j.ecl.2012.03.001.
4. European Medicines Agency (EMA). Batches of the insulin medicine NovoMix 30 FlexPen and Penfill to be recalled, (EMA/ 657469/ 2013). 2013. http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2013/10/WC500153147.pdf .Accessed 09 Feb 2018.
5. Palmer E. Manufacturing glitch leaves novo insulin cartridges with bad fills. 2013.https://www.fiercepharma.com/supply-chain/manufacturing-glitch-leaves-novo-insulin-cartridges-bad-fills. Accessed 08 June 2018.
6. USP 38-NF 33. Insulin Aspart Injection. Rockville, MD: The United States Pharmacopeial Convention, Inc.; 2015.
7. Downing NS, Shah ND, Aminawung JA, Pease AM, Zeitoun JD, Krumholz HM, et al. Postmarket safety events among novel therapeutics approved by the US Food and Drug Administration between 2001 and 2010. JAMA. 2017;317:1854–63.
8. Ebbers HC, NFd T, Hoefnagel MC, Nibbeling R, Mantel-Teeuwisse AK. Characteristics of product recalls of biopharmaceuticals and small-molecule drugs in the USA. Drug Discov Today. 2016;21:536–9. https://doi.org/10.1016/j.drudis.2015.10.020.
9. U.S. Food and Drug Administration (FDA). Recall (Biologics. https://www.fda.gov/BiologicsBloodVaccines/SafetyAvailability/Recalls/ . Accessed 12 Oct 2018.
10. Yu LX, Kopcha M. The future of pharmaceutical quality and the path to get there. Int J Pharm. 2017;528:354–9. https://doi.org/10.1016/j.ijpharm.2017.06.039.
11. Junker B, Kosinski M, Geer D, Mahajan R, Chartrain M, Meyer B, et al. Design-for-six-sigma for development of a bioprocess quality-by-design framework. PDA J Pharm Sci Technol. 2011;65:254–86. https://doi.org/10.5731/pdajpst.2011.00739.
12. Silva Elipe MV, Li L, Nagapudi K, Kook AM, Cobas C, Iglesias I, et al. Application of 19F time-domain NMR to measure content in fluorine-containing drug products. Magn Reson Chem. 2016;54:531–8. https://doi.org/10.1002/mrc.4223.
13. Corver J, Guthausen G, Kamlowski A. In-line non-contact check-weighing (NCCW) with nuclear magnetic resonance (NMR) presents new opportunities and challenges in process control. Pharm Eng. 2005;25:18–31.
14. NovoLog® Mix 70/30 (Insulin aspart protamine and insulin aspart injectable suspension) [package insert]. Plainsboro, NJ: NovoNordisk, Inc.; 2001.
15. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. Rev Sci Instrum. 1958;29:688–91. https://doi.org/10.1063/1.1716296.
16. Yu YB, Feng Y, Taraban M. Water proton NMR for noninvasive chemical analysis and drug product inspection. Am Pharm Rev. 2017;20:34–8.
17. Mathonet S, Mahler H, Esswein ST, Mazaheri M, Cash PW, Wuchner K, et al. A biopharmaceutical industry perspective on the control of visible particles in biotechnology derived injectable drug products. PDA J Pharm Sci Technol. 2016;70:392–408. https://doi.org/10.5731/pdajpst.2015.006189.
18. Briggs KT, Taraban MB, Yu YB. Water proton NMR detection of amide hydrolysis and diglycine dimerization. Chem Commun. 2018;54:7003–6. https://doi.org/10.1039/C8CC03935F.
19. Popel AS. Hydrodynamics of suspensions. Fluid Dyn. 1969;4:14–18. https://doi.org/10.1007/BF01094677.
20. Butt H. Controlling the flow of suspensions. Science. 2011;331:868–9. https://doi.org/10.1126/science.1201543.
21. Shieu W, Lamar D, Stauch OB, Maa Y. Filling of high-concentration monoclonal antibody formulations: investigating underlying mechanisms that affect precision of low-volume fill by peristaltic pump. PDA J Pharm Sci Technol. 2016;70:143–56.
22. Hanslip S, Desai KG, Palmer M, Bell S, Schofield P, Varma P, et al. Syringe filling of high-concentration mAb formulation: slow suck-back pump speed prevented filling needle clogging. J Pharm Sci. 2017;106:3651–3. https://doi.org/10.1016/j.xphs.2017.08.005.
23. Hanslip S, Desai KG, Palmer M, Kemp I, Bell S, Schofield P, et al. Syringe filling of a high-concentration mAb formulation: experimental, theoretical, and computational evaluation of filling process parameters that influence the propensity for

filling needle clogging. J Pharm Sci. 2019;108:1130–8. https://doi.org/10.1016/j.xphs.2018.10.031.
24. Ventola CL. The drug shortage crisis in the United States: causes, impact, and management strategies. Pharm Ther. 2011;36:740–57.
25. Wilson E, Zhu C, Galea S, Marko A, Victoria Urdaneta V, Straus W. Turning up the heat: effect of new vaccine for children's (VFC) program recommendations for use of temperature monitors upon incorrect product storage adverse event reporting. Vaccine. 2018;36:1516–20. https://doi.org/10.1016/j.vaccine.2017.10.059.
26. Rosenberg AS, Verthelyi D, Cherney BW. Managing uncertainty: a perspective on risk pertaining to product quality attributes as they bear on immunogenicity of therapeutic proteins. J Pharm Sci. 2012;101:3560–7. https://doi.org/10.1002/jps.23244.