Bar Codes Decoded

Bar codes are an integral part of modern business in virtually every industry across the globe, providing an automated means of encoding and decoding data for a wide variety of applications. Here, an in-depth breakdown is provided of the past, present and future of this essential technology

Patrick Kearney at Innovatum

Bar codes provide extensive benefits to businesses and consumers alike in terms of efficiency, data accuracy and product safety. Like any business technology, however, bar codes require effective implementation in order to maximise their benefits. This can be a significant challenge in any industry, given the dizzying array of terminology, standards and associated technology (such as software, printers, scanners and so on). Add regulatory compliance to the equation, and the task becomes even more complex for pharmaceutical manufacturers. This article provides a practical summary of bar code standards and regulations applicable to EMEA and FDA regulated pharmaceutical firms, as well as recommendations for effective implementation.

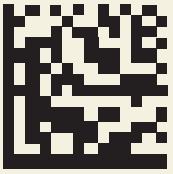
Understanding Bar codes

The seeds for what would ultimately become the modern bar code were sown more than 75 years ago, when JT Kermode *et al* filed a US patent in 1934 for a card sorting machine that used a series of parallel lines for identification and differentiation (1). While Kermode's invention never achieved broad commercial acceptance, it effectively launched the concept of automatic identification and data capture (AIDC), and inspired a multitude of subsequent inventions and technologies in the decades that followed. However, most experts would agree that the birth of modern bar codes occurred some 40 years later with the advent – and standardisation – of the Universal Product Code (UPC) for the North American retail grocery industry in 1973. The early success of the UPC in terms of increased efficiency, data accuracy and inventory control spawned development, innovation and adoption of bar code technology and standards across multiple industry sectors throughout the remainder of the 20th century and up to the present day.

Now, with a better understanding of the origins and evolution, the practical question becomes, 'what is a bar code?' In essence, a bar code is a series of parallel bars and spaces printed and arranged according to a set of rules that defines how data is encoded (that is, printed) and decoded (scanned/read) with little-to-no human intervention. This 'set of rules' is collectively referred to as the symbology. The other major component of any practical bar code is the data structure, which dictates the content and arrangement of information to be encoded within a given symbology. Data structure standards are based on specifications created by independent standards bodies, industry groups, regulatory agencies or any combination thereof. The distinction between a symbology and a data structure is similar to that between a written language and a document type. That is, a given language encompasses the set of rules in terms of characters, spelling and punctuation; while the type of document, such as a news article, cover letter or product manual, specifies the

Figure 1: GTIN and HIBC examples in linear and 2D

The following examples are based on encoding the fictitious NDC number 1234512340, in which 12345 represents the FDA labeller code; 1234 represents the company product code; and 0 represents the company package code.



GTIN with NDC in 2D data matrix symbology

GTIN with NDC in 1D GS1-128 symbology

HIBC with NDC in 2D data matrix symbology

HIBC with NDC in 1D Code 128 symbology

content and organisation of information to be written in a given language.

Today, the most common bar code symbologies found in the pharmaceutical industry include Code 128, Code 39 (also know as Code 3 of 9) and – to a lesser extent – Codabar. All three are classified as linear or 1D symbologies, which means their encoding schemes are based on parallel bars and spaces printed vertically and scanned horizontally. By contrast, 2D matrix symbologies encode and decode data both horizontally and vertically, which allows them to encode significantly more data in a much smaller space than linear symbologies. The most common 2D matrix symbology in pharmaceuticals is data matrix. In terms of the relevant data structures applicable to these symbologies in the pharmaceutical industry, the most common include GS1-GTIN, HIBC and NDC. The following section provides an overview of the organisations from which these data structures originated, and their respective roles in terms of ongoing management and administration.

Relevant Standards Bodies and Regulatory Agencies

As noted previously, a bar code is comprised of both a symbology – or machine-readable language – and a data structure that defines the content and organisation of the data to be encoded. Most industries employ data structures developed and administered by independent standards bodies based on input from key stakeholders within a given industry. Today, GS1 is by far the most prominent standards body for the application of AIDC technology in global trade; encompassing the now-former Uniform Code Council (UCC) and European Article Numbering Association (EAN). The UCC developed and administered the original Universal Product Code (UPC) system in North America in 1973. Four years later, the EAN was formed to serve a similar purpose in the European market. In the decades that followed, the two organisations co-managed the EAN-UCC standards in their respective regions until they officially joined forces under the name GS1 in 2005 (2).

While GS1 has been the primary source of bar code and data structure standards in most industries for decades, it has played a relatively minor role in prescription drugs until recent years. This is mainly due to the early focus on point-of-sale (POS) applications of bar code technology in retail markets, which were largely irrelevant to the unique sales channels employed in the prescription drug industry. Moreover, the POS-centric nature of the early EAN-UCC standards could not effectively address the supply chain challenges associated with tightly-regulated products with a potentially profound impact on public health and safety. Consequently, the Health Industry Business Communications Council (HIBCC) was established in 1983 from a consortium of healthcare trade associations to develop an approach to labelling that could address these unique requirements, ultimately resulting in the Health Industry Bar Code standard (HIBC) (3). Over the years, however, the GS1 and HIBC standards have evolved to the extent that they are now virtually interchangeable in the pharmaceutical industry.

Historically, regulatory agencies have played a significantly less active role in the development of bar code standards and data structures. In 2004, the US FDA published its final rule, titled 'Bar Code Label Requirement for Human Drug Products and Biological Products' (4). The rule added Part 201.25 to Title 21 of the US Code of Federal Regulations (21 CFR Part 201.25), specifying that certain categories of prescription and over-the-counter human drugs manufactured and/or distributed in the US, "…must have a bar code that contains, at a minimum, the appropriate National Drug Code (NDC) number in a linear bar code that meets European Article Number/ Uniform Code Council (EAN.UCC) [now GS1] or Health Industry Business Communications Council (HIBCC) standards" (5).

In the European Union, member states individually administer bar code standards and data structures applicable to pharmaceuticals within their respective borders. The absence of a single, pan-European standard has contributed to a highly fragmented regulatory landscape across the EU: 17 countries employ the current GS1 global standard; 11 countries use the pre-GS1 European standard (EAN); and four countries developed their own standards compatible with neither GS1 nor EAN (6). The resulting challenges are numerous in terms of efficient inter-state trade and the threat to patient safety from the increased risk of counterfeit and/or diverted drugs. In efforts to address these and other challenges in the European pharmaceutical industry, the European Commission (EC) proposed a series of measures in 2008 collectively referred to as 'The Pharmaceutical Package'.

In response to the EC's public consultation on the 'falsified medicines' portion of The Pharmaceutical Package, the European Federation of Pharmaceutical Industries and Associations (EFPIA) recommended a harmonised standard among EU

"As noted previously, a bar code is comprised of both a symbology – or machine-readable language – and a data structure that defines the content and organisation of the data to be encoded"

Table 1: Comparison of bar code data structures for pharmaceuticals

Attributes	Data structures		Implications
	Global trade item number (GTIN)	Health industry bar code (HIBC)	
Length	14 digits, not including potential AI numbers, or parentheses displayed in the human-readable portion of a printed barcode	8 to 25 characters, not including potential AI/DI numbers/characters	Minimal
Governing body	GS1	HIBCC	Both GS1 and HIBCC have a global reach, but GS1 is the larger of the two organisations. GS1 also appears to be more prevalent than HIBCC outside of the North American market
Company ID	Company prefix assigned by GS1 and comprised of two or three-digit geographical code and four to eight-digit company code. NDC labeller code assigned by FDA may be used in place of GS1 company code	Labeller identification code (LIC) assigned by HIBCC and comprised of one alphabetic character and three digits to uniquely identify the company	Both data structures enable unique identification of company. But HIBC requires companies to use its proprietary LIC, while GTIN allows companies to use either the GS1 company prefix in its entirety, or replace company code portion with the FDA NDC labeller code. Thus, for FDA-regulated companies with NDCs, HIBC requires registration and maintenance of two separate company IDs, while GS1 allows for a single company ID
Product ID	Item reference chosen and managed by company. May use product code portion of FDA NDC number	Product or catalogue number (PCN) chosen and managed by company. One to 18 characters. May use 10-digit FDA NDC number in its entirety	HIBC provides more flexibility for variable-length product IDs, as well as the ability to use alphabetic characters. If a company elects to use its FDA NDC number, however, that flexibility is effectively nullified since the NDC is a fixed 10-digit numeric-only value. GTIN provides the flexibility to use only the relevant product code segment of the NDC, which can aid in consolidating and standardizing under a single product ID format.
Package level indicator	Indicator single digit with a value from zero to eight, chosen and managed by the company. May use package code segment of FDA NDC number	Unit of measure identifier (U/M) single digit with a value from zero to eight, chosen and managed by the company	As GTIN and HIBC are virtually identical in this regard, the implications are minimal. However, since GTIN provides segmentation of the NDC, it may reduce data redundancy and simplify long-term system administration
Supported Symbologies	GS1-128, interleaved 2 of 5, GS1 DataBar, and GS1 data matrix	Code 128, Code 39, Aztec Code, data matrix, MicroPDF417 and QR Code	HIBC supports more 2D symbologies, but the practical advantages are minimal in a pharmaceutical setting
Alpha-numeric	Numeric only	Alpha-numeric	As noted previously, HIBC's ability to encode alpha-numeric data theoretically provides more flexibility. In practice, however, the advantage for FDA-regulated companies is minimal at best given the numeric-only format of an NDC
Supplemental data	Can be appended via any GS1 standard application identifier (AIs)	Can be appended in HIBC secondary bar code via most ISO/IEC 15434 standard data identifiers (DIs), or 2 of the GS1 standard application identifiers (AIs)	Both standards allow for encoding additional product attributes that can aid in batch/lot traceability, product tracking and verification. HIBC also provides some minor cross-compatibility with GS1 standards via AIs

member states based on a GS1 data structure encoded within a 2D data matrix bar code (7). The European Parliament subsequently adopted the proposed falsified medicines legislation in February 2011, which was then published on 1 July, 2011 in the *Official Journal of the European Union*. The new legislation will go into effect 2 January, 2013. Similar to the FDA's rule on bar code label requirements, however, the EU legislation does not require use of a specific data structure. Thus, given the somewhat ambiguous regulatory stance in both the US and European markets, it is important for pharmaceutical companies to understand the various applicable data structures and their respective implications. To that end, the following section provides a brief description of each data structure, followed by a summary of potential considerations for pharmaceutical companies in planning for effective implementation.

Data Structures and Implementation Considerations

The most common data structure of the GS1 system is the Global Trade Item Number (GTIN). A GTIN is comprised of multiple data elements concatenated within a bar code to uniquely identify 'trade items,' which GSI defines as: "…any product or service upon which there is a need to retrieve pre-defined information; this product or service may be priced, ordered, or invoiced at any point in the supply chain. This includes individual items as well as all of their different packaging configurations" (8). GTINs in a pharmaceutical context are generally 14 digits long.

While variations abound, all GTINs include three basic data elements: a company prefix assigned and maintained by GS1 to uniquely identify the company; an item/product reference number assigned and maintained internally by the company; and a check digit, which is the result of an algorithm using the preceding numeric data to calculate an expected value read by the bar code scanner in order to ensure data integrity and reduce scan/read errors. The basic GTIN structure may also be configured to address different packaging levels (unit, case, and pallet). Moreover,

GS1 application identifiers enable standardised encoding of additional product attributes like batch/lot number, expiration date and serial number in a GS1 bar code to aid in unit-level traceability and verification between the manufacture to the ultimate point of dispensation. Applicable bar code symbologies for encoding GTINs include GS1-128 (subset of Code 128), Interleaved two of 5, GS1 DataBar (formerly known as RSS) and GS1 data matrix (a subset of data matrix).

As noted in the previous section, the HIBCC was formed in 1983 to develop a healthcare-specific approach to bar code labelling that could address the unique challenges of the industry for which EAN-UCC (now GS1) specifications of the day were insufficient. The result was the Health Industry Bar Code (HIBC) standard. Similar to the modern GTIN described above, the HIBC data structure is comprised of multiple data elements that uniquely identify the company, product and packaging level, as well as a check character to ensure data integrity and reduce scan/read errors.

In the HIBC standard the company identifier is a four-character labeller identification code (LIC) assigned and administered by the HIBCC. The product identifier is a one- to 18-character product or catalogue number (PCN) assigned and administered by the company. The packaging level is represented by a one-digit unit of measure identifier (U/M) that is also assigned and administered by the company. Another similarity to GS1-GTIN is the ability to encode additional product attributes like batch/lot number, expiration date and serial number. Specifically, the HIBC secondary data structure is compatible with most ISO/IEC 15434 Data Identifier specifications. It is also compatible with two of the GS1 Application Identifier specifications. Applicable linear bar code symbologies include Code 128 and Code 39, while 2D symbologies include Aztec Code, data matrix, MicroPDF417 and QR Code.

The FDA's National Drug Code (NDC) data structure, when used in accordance with 21 CFR Part 201.25, is a 10-digit number comprised of three distinct segments. The first segment is a four or five digit labeller code assigned by the FDA to uniquely identify a pharmaceutical manufacturer or distributor. The second segment is a three or four digit product code assigned by the manufacturer/distributor to identify the strength, dosage form and formulation of a given drug product. The third segment is a one or two digit package code assigned by the manufacturer/distributor to identify the commercial package level (unit, case, pallet and so on). Thus, the NDC structure is very similar to both the GTIN and HIBC data structures described above. As noted previously, however, the FDA provides no detailed specifications in terms of bar code symbology: 21 CFR Part 201.25 merely states that bar codes should conform to either GS1 or HIBCC standards.

Collectively, the preceding descriptions clearly illustrate significant commonalities among the GTIN, HIBC and NDC data structures. Now the question is whether to utilise GTIN, HIBC or some combination thereof, since both are capable of encoding an NDC in accordance with FDA 21 CFR Part 201.25. To that end, Table 1 (see page 48) provides a side-by-side comparison of the GTIN and HIBC standards with pertinent considerations in the selection process.

HIBC has certain theoretical advantages in terms of alpha-numeric flexibility and the capacity to incorporate NDCs and certain GS1 application identifiers for encoding additional product attributes. It also supports more linear and 2D symbologies. However, that flexibility comes at the expense of simplicity relative to GS1 standards. In addition, GS1 seems to have a greater world-wide adoption, and is at the forefront of emerging standards and legislation with regard to serialisation and ePedigree initiatives identified as a top priority by state and federal government in the US, as well as the EU. Consequently, GS1 would appear to be the optimal choice for companies that are not otherwise constrained by local regulations and/or requirements of key business partners.

References

1. Palmer RC, *The Bar Code Book*, 1995
2. *GS1 Timeline*, http://www.gs1.org/about/media_centre/timeline
3. *Why Use the HIBC? – FAQs*, http://hibcc.org/AUTOIDUPN/standards.htm
4. *Federal Register*, February 26, 2004 (69 FR 9120).
5. US Code of Federal Regulations Title 21, Part 201, Subpart A, Section 201.25, Paragraph C
6. *Verification of Pharmaceutical Products at the Point of Dispense*, European Federation of Pharmaceutical Industries and Associations, http://www.efpia.eu/content/default.asp?PageID=559&DocID=8773
7. *EFPIA Response To The European Commission Public Consultation in preparation of a legal proposal to combat counterfeit medicines for human use*, http://ec.europa.eu/health/files/counterf_par_trade/doc_publ_consult_200803/114_efpia_en.pdf
8. *An Introduction to the Global Trade Item Number* (GTIN), GS1 US

About the author

Patrick Kearney is Senior Consultant – Regulated Labelling at Innovatum. Patrick has worked in the enterprise label software industry since 1998. He currently heads Pre-Implementation Consulting Services at Innovatum – an enterprise compliance software and services company dedicated to the IT challenges of regulatory compliance in the life sciences. In this capacity, he has helped numerous international pharmaceutical, medical device and biologics manufacturers meet FDA and EMEA regulations. Prior to joining Innovatum in 2004, Patrick spent six years managing Inside Sales and Major Accounts at Seagull Scientific – a leading developer of barcode and RFID label design and printing software. He successfully completed Ohio University's AIDC Technical Institute programme and holds a BS in Organisational Communication from the University of Portland. **Email: patrick@innovatum.com**