

Two Dimensional Bar coding with PDF-417



By Robert Burrell
© 2009 PrimeCare Systems, Inc.



Two Dimensional Bar coding with PDF-417



By Robert Burrell

© 2009 PrimeCare Systems, Inc.

Abstract

Two-dimensional bar coding systems have recently sprouted applications in many industries as a result of the need for a small, secure bar code that can store more data than traditional one-dimensional bar codes (e.g. UPC's). The PDF-417 technology provides for these requirements, particularly with its features such as data compaction and error correction. PrimeCare's Patient Management System software simplifies many management and inventory tasks in the health care industry using bar coding technologies, and with these 2-D technologies PrimeCare's products can also generate bar codes that are largely damage resistant.

Introduction

For over 30 years, bar coding has been a technique used in the manufacturing industry for shipping and receiving as well as inventory purposes. Bar codes have traditionally offered a key-less entry point into databases for quicker, more efficient management systems. Today, bar coding has become popular in many commercial arenas such as airlines, retail stores, security applications, and postal services. As the technology has become more and more prevalent, there has been an increasing desire to store more data within a bar code, turning it into a small data source as opposed to a unique key into a larger database. Also, for many applications, bar codes may be subject to wear-and-tear or otherwise harmful damage that jeopardize the readability of the bar code. These factors led to the development of many styles of two-dimensional bar coding systems that offer a much larger space for data and more reliable error correction techniques among other features.

Traditional (one dimensional) bar codes store information horizontally, while the vertical aspect to the symbol provides a small degree of redundancy. Taller 1-D bar codes have more vertical area that can be damaged or removed and still remain readable by the scanner; the scanner only needs one clear horizontal path in order to have a successful read. Accordingly, two dimensional bar codes store information both horizontally and vertically. Since most 2-D applications do not face strict space requirements, the bar code is able to encode redundancy in both directions making the image more damage resistant.

The biggest advantage of using a 2-dimensional bar code is the larger data capacity. One-dimensional bar code styles are typically restricted to a small character set; standard Code 39 encoding, for example, is limited to a 44 character set - single case letters, numbers, and selected punctuation. Expansion to a larger character set (Full-ASCII Code 39) that includes cased letters and additional punctuation increases the average width of the bar codes by about 25%.

Figure 1 illustrates printed size differences between typical 1-D and 2-D bar codes, as well as the larger data capacity of 2-D versions.



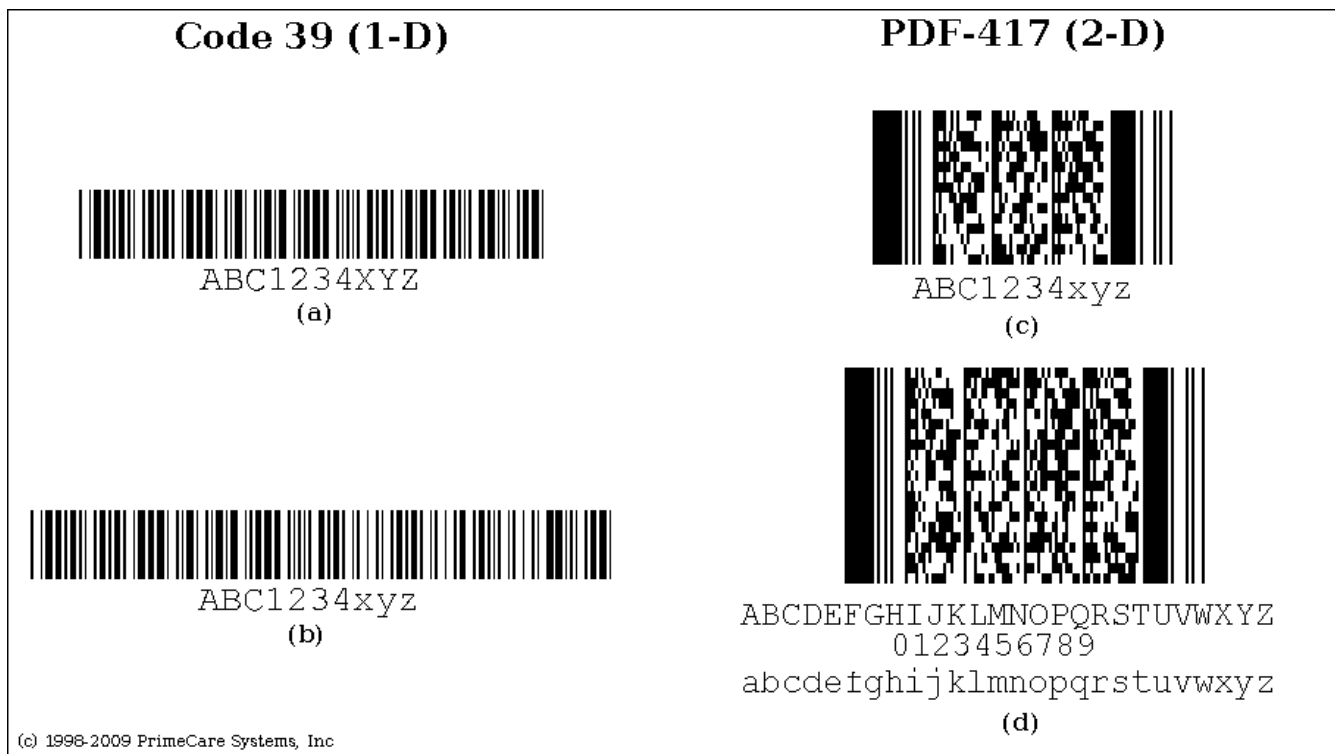


Figure 1. Code 39 vs. PDF-417.

A simple string is represented by the bar code in (a). Alpha-characters are single cased unless the full-ASCII expansion encoding is used such as in (b). The simple string is then given in 2-D format in (c). The size of the bar code does not significantly increase for the much larger string represented in (d).

PDF-417, the Emerging Standard

There are a large number of 2-D encoding styles for bar codes, many designed with specific applications in mind. Figure 2 shows some of the most often used bar coding formats.






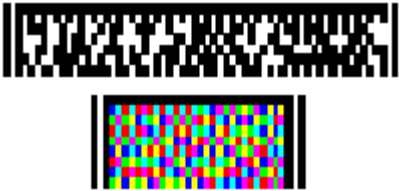
 <p data-bbox="396 594 565 621">Aztec Code</p>	<p data-bbox="816 163 1464 380"><i>Designed for ease of printing and ease of decoding; data blocks are arranged on a square grid to compose a squarely shaped symbol, characterized by its centered bullseye finder. Aztec code symbols are used in such applications as airline and railway ticketing.</i></p>
 <p data-bbox="383 831 568 858">Data Matrix</p>	<p data-bbox="816 674 1445 852"><i>Designed to pack information into a particularly small area; can store up to 2335 characters, and is scalable (and readable) down to a couple square millimeters, making it a useful format for small applications.</i></p>
 <p data-bbox="418 1138 545 1165">QR Code</p>	<p data-bbox="816 966 1450 1220"><i>(Quick Response Code) square matrix code with finder patterns in three corners, limited to 4464 alpha numeric characters. One important feature is its ability to directly encode Japanese characters. QR Code was designed for rapid reading using CCD array cameras and image processing technologies.</i></p>
 <p data-bbox="402 1455 555 1482">UltraCode</p>	<p data-bbox="816 1236 1456 1491"><i>Features advanced language/code page handling provisions, error correction, both black & white and color versions. Different from most 2D bar codes in that they are long and narrow (similar to 1-D bar codes) and therefore are useful for applications with more narrow, linear space requirements.</i></p>

Figure 2. Two-Dimensional bar code type examples.

The most prominent 2-D type, PDF-417, has proved to be as the most efficient in packaging the most amount of data into the smallest space, while providing many other useful features. Invented by Ynjiun Wang in 1991 at Symbol Technologies (now Motorola), the "Portable Data File" (PDF) symbology is essentially a stacked version of the classic horizontal symbol; Figure 3 illustrates the symbol's structure. According to the official specification document, "up to 1100 bytes or 1800 ASCII characters can be encoded in a PDF417 symbol," although in practice bar codes are usually generated with upper



bounds anywhere between 300 and 800 characters. Similar to its 1-D predecessor, these limitations mainly stem from the capabilities of typical scanners.

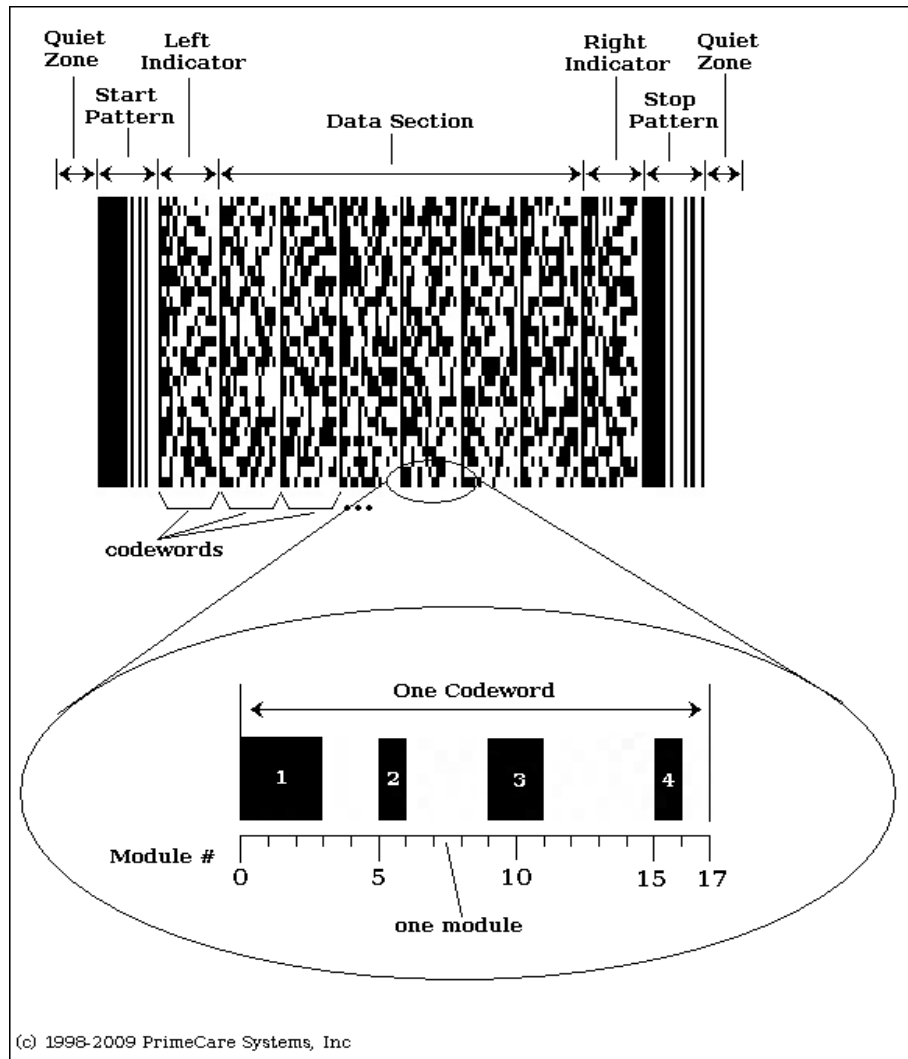


Figure 3. PDF-417 Structure.

The PDF-417 bar code is made up of between 3 and 90 rows, with each row containing a series of grouped sections called codewords (the number of columns varies as it is configurable by the user). Each codeword consists of 17 modules that each contain 4 bars and spaces (hence, 417).

Along with the increased data space and reduced print size, the PDF-417 symbology offers other useful features. For instance, the technology allows for user-specified print sizes, meaning the user can print shorter, wider bar codes or taller, slimmer ones. The overall size of the bar code, however, is determined by a number of factors such as the amount and type of data, compaction mode selections, error correction settings (see below), etc. Again, the height-to-width ratio of the bar code is usually determined by the abilities of the scanner.

Data compaction is another unique feature that allows data to be more tightly packed for a smaller print



size. PDF-417 offers three compaction modes: text mode, which operates on printable alpha-numeric characters (ASCII 9,10,13,32-127); binary mode, which operates on 8-bit binary data (ASCII 0-255); and numeric, which allows only numeric values to be encoded.

Another feature is the optional extension to MacroPDF417, which allows multiple bar codes to be linked together for one larger set of data. Each Macro bar code contains an identifier and sequence number so that a number of codes can be scanned in or out of sequence and packaged together.

PDF-417 and Error Correction

One important aspect of bar coding has always been providing error correction mechanisms and some level of data redundancy. Slight glitches can occur in bar code printing could render incorrect data, and damage to printed bar codes could render them unreadable. As mentioned above, one-dimensional codes provide some data redundancy in the height of the symbol; tall bar codes can usually be torn in half horizontally and remain readable. To provide for error correction, 1-D codes have traditionally used a check digit system that verifies the data when read.

The larger space capacity for 2-D symbols provides for additional data to be stored in the bar code that achieves both redundancy and error correction. Instead of check digits, PDF-417 uses a Reed Solomon error correction technique that is similar to that used in multimedia applications (CD's, DVD's and Blu-Ray Discs), telecommunications and transmission technologies (e.g. DSL), and broadcasting (Digital Video Broadcasting). The correction technique works by oversampling a polynomial constructed from the data; the receiver (scanner), then, only has to be able to read a majority of the data "samples" in order to be able to reconstruct the original data.

PDF-417 encoding provides 9 levels of error correction (0-8), which increasingly allow more and more of the bar code to be damaged without becoming unreadable. There is, however, a linear relationship between the level of error correction and printed size of the bar code, so the setting should be chosen carefully. Accordingly, the ideal level selection usually will depend on the amount of data to be encoded and the desired final size of the bar code. Table 1 shows the number of codewords and data bytes that are added to the packaged symbol according to the error correction level selection.

EC Level	0	1	2	3	4	5	6	7	8
EC Codewords Generated	2	4	6	8	16	32	64	128	512
Data Codewords			1 – 40	41 – 160	161 – 320	321 - 863			
Data Bytes Encoded			1 – 56	57 – 192	193 – 384	385 – 1035			

Table 1. Codeword addition for PDF-417 Error Code levels.

PDF-417 Applications in PrimeCare's Patient Management System

Two-dimensional bar coding has many applications in the health care industry, particularly for administration, management, and inventory purposes. PrimeCare's flagship product, the Patient



Management System, Version Nine Client (V9 Client), takes advantage of the many uses for bar coding, particularly using the PDF-417 format.

One of the most used bar code applications lies in patient appointment ticketing. At the time that a patient's appointment is scheduled, the V9 Client automatically prints an appointment ticket, which contains some basic readable information about the appointment as well as a PDF-417 bar code. Packed into a small area, this bar code can contain additional basic information about the patient information, as well as other data items such as the appointment date, time, location, and the scheduled physician. Also included in the bar code are links to database items such as the patient's electronic medical record. At a later time, the patient's ticket can be easily scanned to initiate V9 Client tasks such as rescheduling or changing the appointment, checking in the patient for a visit, or loading the patient's medical records.

The 2-D bar coding system also has other V9 Client applications, such as pharmacy inventory and inpatient bed management. For more information about PrimeCare's Version Nine Client and its bar coding applications, please visit www.versionnine.com.

Testing Demonstrations

Multiple test situations were used to analyze the redundancy and error correction factors for PDF-417 bar codes used by PrimeCare's Patient Management System appointment ticket printing process. The figures below illustrate typical test results.



Original

Crumple Test



PASS

PASS

(c) 1998-2009 PrimeCare Systems, Inc.

Figure 4. Original / Crumple Test. An original, unaltered version of the V9 Appointment ticket is shown left, and a crumpled test version is shown right. Both successfully scanned.



Horizontal Test

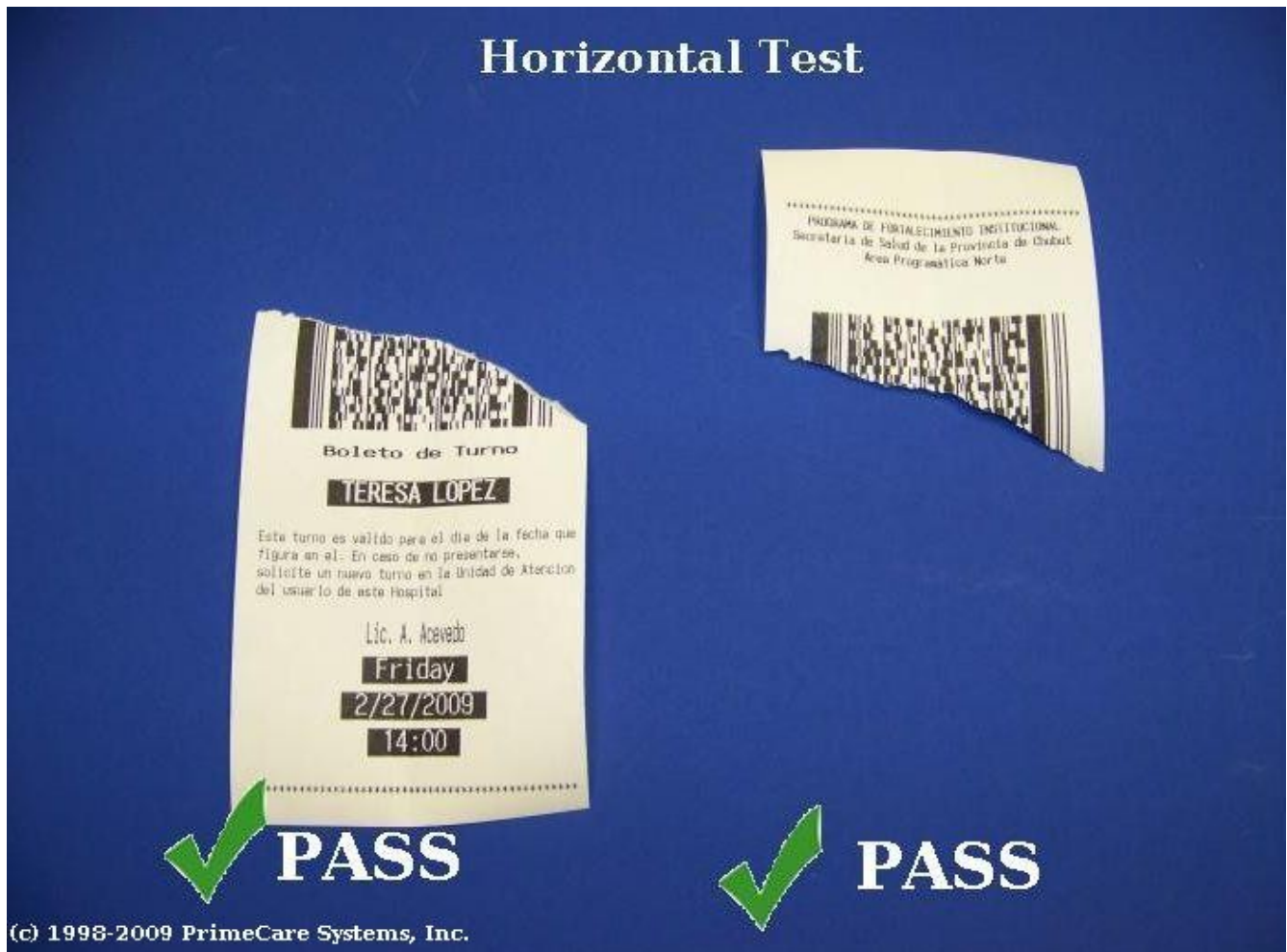


Figure 5. Horizontal Test. *Another ticket was torn in approximately in half horizontally. Surprisingly, both ends of the torn ticket successfully scanned.*



Vertical Test



(c) 1998-2009 PrimeCare Systems, Inc.

Figure 6. Vertical Test. *The next ticket was torn vertically. The larger half successfully scanned while the narrow half did not.*



Central Removal Test



PASS



FAIL

(c) 1998-2009 PrimeCare Systems, Inc.

Figure 7. Central Removal Test. *A large section was removed from a ticket. The remaining piece still scanned, while the removed piece did not.*



Write-On Test

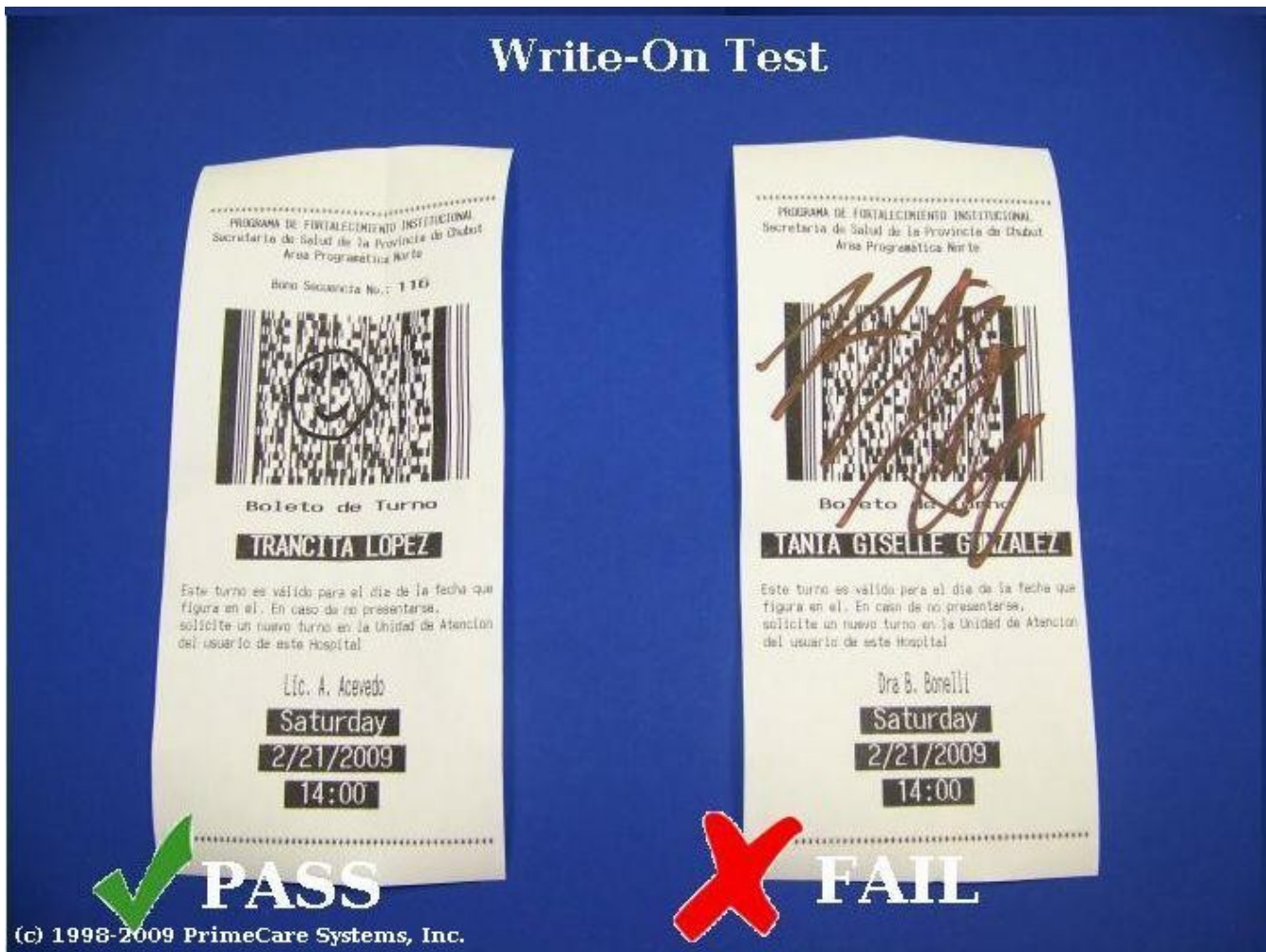


Figure 8. Write-On Test. Two tickets were written on. A small amount of writing (smiley face, left) still rendered the bar code readable, while a large amount (right) did not.

About PrimeCare Systems, Inc., and the Version Nine System

The PrimeCare Patient Management System, Version Nine (the Version Nine System) is an electronic medical records system developed by PrimeCare Systems, Inc., that incorporates lessons learned through years of clinic practice and experience to the documentation of the physician-patient encounter. Centered around more than 280 problem-specific, medically appropriate patient questionnaires, the system is an intuitive, user-friendly interactive system for physicians, staff and patients. The system is an internet-based system that requires of the physician's office only inexpensive commercial off-the-shelf computers and a broad-band internet connection to be put into immediate use.

For more detailed information or to download an evaluation version of the PrimeCare System, please visit the Version Nine dedicated web site, <http://www.versionnine.com>.

Robert Burrell is a software engineer at PrimeCare Systems, Inc.

