Pestilential Protocol: How Unsecure HL7 Messages Threaten Patient Lives

Christian Dameff MD[1], Maxwell Bland[1], Kirill Levchenko PhD[1], Jeff Tully MD[2]

Abstract:

This paper describes the ubiquitous role of the HL7 standards in the transfer of patient data and facilitation of the life-saving clinical workflow within hospital networks, particularly with respect to orders placed by physicians and other healthcare providers to gather crucial information to guide clinical decision making. These HL7 standards are often implemented in an unsecure fashion resulting in unauthenticated, unvalidated, plaintext transmission of sensitive data across flat networks, leaving systems vulnerable to a variety of attacks. In addition to the significant privacy implications such deployments raise, the corruption of data integrity resulting from an attacker may ultimately lead to patient morbidity and mortality.

We developed a tool to perform an automated man-in-the-middle attack exploiting the lack of encryption and authentication in most HL7 implementations to maliciously change laboratory results from normal values to those consistent with serious illness. These “pathologic payloads” may result in clinical staff mistakenly believing a patient has a particular disease requiring a treatment that can ultimately harm the patient.

This project aims to explore concepts which challenge the implicit trust placed in medical technology infrastructure by care providers, and contributes to the ever-growing list of compelling reasons for the design and implementation of secure medical devices, protocols and networks throughout healthcare delivery organizations.

Keywords: Healthcare, HL7, protocols, vulnerabilities, electronic medical record, laboratory system, encryption.

Introduction:

The practice of medicine is one built upon the foundation of hypothesis generation and testing. Disease presents in a constellation of subjective symptoms and physical exam findings, and medical professionals are trained to use pattern recognition, differential diagnosis, and clinical algorithms to organize a wealth of data

[1] University of California, San Diego
[2] University of California, Davis Medical Center

into actionable evidence. In today's modern healthcare system, clinicians are supported by a vast technological infrastructure that includes electronic medical records, automated ordering systems, digital imaging, and internet connected medical devices.

From the earliest days of medical and nursing school, students are trained in the use of these technologies to care for patients, quickly developing a dependence on these tools to streamline an ever more complex and increasing workload (1). Paper charts, prescription pads, and x-ray illuminators are often more foreign to the modern clinician than a rare disease or experimental medicine.

The myriad infrastructure and bedside technologies facilitating healthcare delivery are sustained by an equally large number of support technologies. One of the most important is the Health Level 7 (HL7) standard- a near ubiquitous communication framework that dictates the protocol for almost all clinical workflows including: sharing of electronic health information, ordering and viewing of laboratory test and diagnostic imaging, hospital admissions, and medication and therapy orders (2). The culmination of a standards development process beginning in academia in the late 1970s and continuing until widespread adoption in the 1990s, HL7 served as a tool to interface between the disparate proprietary systems hospitals were using to manage laboratory results, computerized order entry, pharmacy dispensing, and dozens of other clinical tasks (3).

Today's HL7 standards are typically implemented either as version 2 (designed in 1989, non-XML based, and by far the most common implementation in healthcare today) and version 3 (first implemented in 2005, XML based, and less widely adopted in part due to lack of backward compatibility). A typical HL7 message (Figure 1) is transmitted in a plain text, pipe-delimited syntax. Messages typically flow bidirectionally between a laboratory information system (LIS) and the medical record system (MRS); laboratory test orders are transmitted to the LIS, and laboratory test results are transmitted to the MRS. Each transmission typically involves the establishment of a TCP client-to-server connection and connections are often short-lived, therefore vulnerable to ARP-spoofing attacks in non-static MAC address environments.

```
MSH|^~\&|Rapidcomm |Hospital|OpenEMR|Hospital|20180719164041||ORU^R01|0C0AGPD228ZGM001D808|P|2.4|||AL|AL|
PID|||99||TTT^BT|||U|
ORC|RE|
OBR|1|6|0C0AGPD228ZGM001D808|666^Venous Blood Gas|R|||||O||||BLDA^^^^^^P|^Administrator|||||||||F|
OBX|1|ST|pH||7.12||7.350-7.450|L|||F|||20150528093432||||^^07143^RAPIDPoint 405|20150528093432|
OBX|2|ST|pCO2||27|mmHg|35.0-45.0||||F|
OBX|3|ST|pO2||77|mmHg|75.0-100.0|H|||F|
OBX|4|ST|tHb||21.5|g/dL|12.0-18.0|H|||F|
OBX|5|ST|O2Hb||97.0|%|94.0-97.0||||F|
OBX|6|ST|COHb||0.4|%|0.5-1.5|L|||F|
OBX|7|ST|MetHb||0.6|%|0.0-1.5||||F|
OBX|8|ST|HHb||2.0|%|0.0-5.0||||F|
OBX|9|ST|HCO3act||7|mmol/L|||||F|
OBX|10|ST|BE(B)||-12|mmol/L|||||F|
OBX|11|ST|sO2||98.0|%|92.0-98.5||||F|
OBX|12|ST|Samp. Type||BLDV|||||F|
```

[FIGURE 1]

There is no encryption at the level of the protocol, and the HL7 Standards Committee, Health Level Seven International, assumes that encryption happens “below the application layer (4).” Although no accurate data exists reporting the prevalence of HL7 message encryption, many clinical software tools which use the protocol lack out-of-the-box support for encryption.

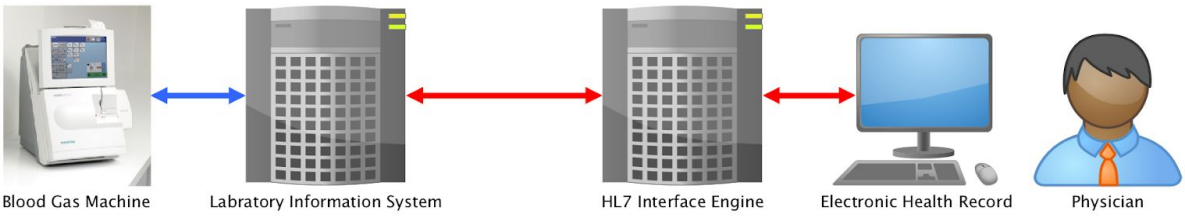
Additionally, HL7 standards lack verification of message source and authentication of message transmission. Interoperating systems thus trust, incorporate, and store clinical data that is vulnerable to malicious alteration. This potential for forgery presents an equal concern to that of passive sniffing; an attacker who has intercepted packets between machines and software communicating over the protocol can readily modify data using regular expressions (regex) and direct string manipulation. While modern protocols often include explicit cryptographic challenges and guarantees, HL7’s standard does nothing more than *suggest* that the messages should be transferred over some form of transport layer security. Since hospital IT staff are often underequipped and undertrained in the operating specifics of network-facing medical software, it is understandably the case that such a suggestion is often not taken. Hospitals are therefore left to rely on VLAN segmentation and other imperfect methods to guarantee the integrity of messages sent over the network. Finally, the systems which use HL7 often lack recent security updates because of their age and the risk to patient safety should an update render the machine unusable. This is a natural result of the long process for regulatory approval of medical devices and the use of these devices in high-risk, time-sensitive environments (emergency rooms).

These qualities render HL7 messages insecure. In the absence of encapsulating encryption and authentication, the protocol poses a prime target for attack. In this paper we describe the creation and implementation of a tool designed to perform a man-in-the-middle (MitM) attack on HL7 message transmission from a laboratory blood

analysis device to a cross-platform HL7 interface engine, intercepting and subsequently modifying the HL7 message to change information that may ultimately result in patient harm by disrupting the clinical decision making process.

Methods (Technological):

A testbed was constructed consisting of an open-source electronic medical record system (OpenEMR v5.0.1, OEMR), a HL7 connection integration engine (Mirth Connect v3.6.1 NextGen Healthcare, Irvine CA), a laboratory information system (Siemens RapidComm v6.1), and a RAPIDPoint 405 Blood Gas Analyzer machine (Siemens Healthcare v3.9). The electronic patient record, HL7 Interface engine, and laboratory information system were hosted on two PCs (Intel NUC Model: 961252) running Ubuntu Linux (Version 18.4 LTS) and Microsoft Windows Server 12 respectively. Network connectivity was achieved using ethernet and an unmanaged switch (Linksys EtherFast 10/100). Static IPs were assigned to each device in the testbed [Figure 2]. This testbed was built to mirror systems deployed in real clinical environments using devices and software that currently care for patients.



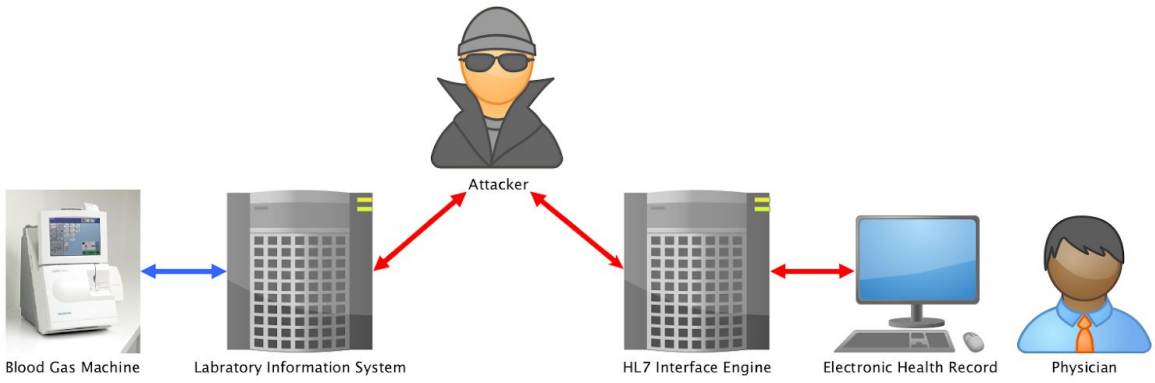
[FIGURE 2]

Our simulated attacker performed a standard man-in-the-middle attack using a tool we developed named Pestilence [Figure 3]. This tool was written in Python using standard packet manipulation (Scapy), threading, and TCP/IP socket libraries. At launch, Pestilence initiates arp-poisoning against targets found through either dynamic (monitors network traffic and determines targets) or static (preconfigured from configuration files) runtime configurations. Since communications between targets using HL7 are in the form of short-lived, intermittent TCP connections, this approach is effective so long as the network infrastructure permits. Once the targets are determined,

the program spawns two threads, one for each victim, with the goal of mediating bidirectional communication of patient health information (PHI).

Each thread is responsible for one direction of the communication. The thread first spawns a TCP server which waits for a connection from it's designated victim; when a TCP client connection is established by the target victim to Pestilence, hereafter referred to as connection A, the tool immediately responds as if it were the *real* server (the non-target victim); simultaneously, Pestilence instantiates a TCP client connection to the non-target victim, hereafter referred to as connection B. It follows that the state machines involved in managing each connection can be handled independently, and the sequence (SEQ) / acknowledgement (ACK) numbers used to validate data transfer are unique to each connection. Data can thereafter be queued from connection A, modified in an arbitrary manner, and passed into connection B without the need to account for any disagreement in SEQ/ACK, as would be the case in a forwarding-based setup.

Using the above method, paired with user defined semantics for which data modification and monitoring, Pestilence can freely modify and read data transmitted between the two victim machines. As HL7 laboratory test orders flow between the MRS and the LIS (victim connections A and B, respectively) through Pestilence, a list of possible target patients and PHI is populated via regex and stored in the software's memory. Once a target is selected, "pathologic payloads" specifically developed to cause iatrogenic harm can be injected into the data queue between the victims; the resulting packet stream will thereafter contain malicious laboratory values, which, once displayed in the MRS, will be used by the clinician in determining treatment.



[FIGURE 3]


PESTILENCE
Version 1.3.3.7
Maxwell Bland, Christian Dameff, Jeff Tully
2018
VICTIM LIST (4)
P381942 | Kirill Levchenko | Urinalysis
P323222 | Nishant Bhaskar | Urinalysis
P348202 | Yifan Li | Blood Gas
P424242 | Christian Dameff | Blood Gas
PATHOLOGIC PAYLOADS (1)
VBG_BAD_THINGS — Created 2018.07.30
├─ pH — 7.19
├─ PCO2 — 29
├─ PO2 — 77
├─ O2 — 95
├─ HCO3 — 16
└─ BASE EXCESS — -8
Cultivating Crops ...
Initializing Server ...
Servers Ready ...
Poisoning ARP Tables ....
Waiting for Victims ...
Ready to go!
~ The art of communication is the language of leadership.
>>> set victim Christian
~ Alright, victim set to Christian Dameff.
~ Their current orders are Blood Gas.
~ A leader is one who knows the way, goes the way, and shows the way.
>>> set payload VBG_BAD_THINGS --times-to-infect=1 --fuzz-values-percent=0.01
~ Jesus, that one looks quite bad, are you sure about this?
~ Okay, whatever, waiting for victim ...
~ Still waiting ...
~ Still waiting ... don't you have anything better to do?
~ Lab results modified! Show pcap [Y/N]?
>>> N
~ A good leader takes a little more than his share of the blame, a little less than his share of the credit.
>>> exit

[FIGURE 4]

Methods (Clinical):

Pathology payloads appropriate for different classes of laboratory devices were constructed using clinical knowledge of corresponding values consistent with various disease presentations. Specific disease states were developed to induce misinformation that would alter clinical decision making and lead to a potentially deadly iatrogenic event. Two such disease states are detailed below.

Diabetic ketoacidosis (DKA) is a serious complication of diabetes mellitus resulting from an inability for the body to metabolize glucose secondary to lack of insulin, and subsequent compensatory oxidation of fatty acids resulting in acidemia and ketosis. Patients can present with this condition with a variety of non-specific complaints, including abdominal pain, increased frequency of urination, visual changes, and vomiting.

Blood gas analysis is a common technique used to assess the acid-base and electrolyte status of sick patients. A venous blood gas sample of a patient in DKA is characterized by low pH and calculated bicarbonate, and high blood glucose and anion gap. Clinicians provided such a sample in a patient presenting with one or more of the aforementioned complaints would likely consider DKA within their differential diagnosis.

Thus, a patient presenting with a relatively benign complaint like viral gastroenteritis or dehydration could, as part of a laboratory work up, have a venous blood gas drawn which, using Pestilence's specific payload, would be corrupted to suggest DKA. The treatment for DKA includes intravenous fluid replacement and the initiation of an insulin infusion, which in a patient NOT in DKA would result in a catastrophic crash in blood glucose levels potentially leading to life threatening seizures, coma, and cardiac arrest.

The basic metabolic panel (BMP) is a further example of a commonplace, low-cost, high yield laboratory study that could be corrupted by a Pestilence pathologic payload. Comprising a survey of the body's main ions as well as indicators of kidney and liver function, the BMP is among the frontline tests ordered whenever a clinician suspects any abnormalities of electrolytes. A wide variety of medications taken for conditions such as high blood pressure, kidney disease, or adrenal insufficiency have effects on the body's balance of electrolytes, and as such, electrolyte imbalances are a very common presentation in the clinic or hospital ward. Manipulation of data points along this spectrum could result in a host of similarly devastating consequences- the patient whose potassium is changed from a normal value to a seemingly extremely low value- and who is also on a diuretic pill that could conceivably explain such a value while the patient was still symptomatic- could receive an intravenous dose of potassium intended to treat the falsely low report but instead having been given essentially a lethal injection.

Discussion:

We will now attempt to frame the weaknesses of HL7 systems within the broader context of healthcare security. So far, we have demonstrated the technical ease of performing a TCP man-in-the-middle attack on HL7 communications and modifying data with the intention of disrupting the clinical decision making process. While the security challenges of the HL7 protocol have been previously discussed (5), this project is novel in that it combines a technical exploitation of HL7 traffic with clinical expertise. The result of this collaboration is a sophisticated attack with the potential for destructive mutations of the patient care process, and is of chief concern to those patients which may be considered “high risk”, i.e. heads of state, celebrities.

Discussion of cybersecurity in the healthcare space often focuses heavily on regulatory compliance and the protection of patient privacy and data. The “CIA triad” is a common data security paradigm and focuses on ensuring the confidentiality, accessibility, and integrity of data within a system. To date, the majority of attention has been placed on confidentiality, and understandably so, as healthcare records often contain valuable personal information -- from social security numbers to financial records and embarrassing medical conditions.

Additionally, recent ransomware attacks affecting hospitals have drawn notice to the issues that arise in attacks on data accessibility- with electronic medical records shuttered and pharmacy, laboratory, and admitting systems offline, patient care grinds to a halt. Conventional wisdom would suggest that redirecting a finite number of resources within a healthcare delivery organization to manage and deal with accessibility and confidentiality attacks would limit the ability of the organization to use the same resources towards other objectives, including direct patient care. Indeed, researchers have already begun to report epidemiologic data suggesting that relatively simple data breaches are associated with increased morbidity and mortality secondary to declines in quality of care as a result of the high costs of remediation and security (6).

In comparison, data integrity attacks have received less attention in healthcare cybersecurity. Healthcare is distinct from other sectors in that the manipulation of critical infrastructure has potential for direct impact to human life, whether this be through direct manipulation of the devices themselves or through the networks which connect them. With the almost ubiquitous integration of networked technologies into the clinical workflow, medical decision making depends upon the integrity of data within these systems, and more generally, the integrity of the systems themselves. *Any successful attack on a medical subsystem will not only disrupt the ability of clinical professionals to provide care, but may also result in serious patient harm.* The integrity attack outlined in this paper is just one of the many possible given the current vulnerable surface area of

hospital technological infrastructure. It follows that HL7, while pestilential in isolation, is a symptom of a more fundamental problem in contemporary healthcare IT; the protocol's prevalence demonstrates the flawed, legacy-device foundations of the current patient care environment.

This hypothesis was supported during our analysis of the machines and software involved in patient laboratory analysis; a port scan of the embedded healthcare machines used in the testbed found them using legacy operating systems with services vulnerable to remote exploit. After gaining persistence on these machines through these services, we had the capability of mounting the pathology payload injection attack presented in the technological methods section without the need for arp poisoning, MITM, or network activity beyond an minor initialization-stage perturbation. These machines are not capable of being upgraded to remove the security vulnerabilities, and were exploitable via pre-existent metasploit modules. Given that such issues would show up using any consumer-level vulnerability scanner, internet-facing hospital infrastructure is ripe for worms, contains infinite-day exploits, and proves a fine target for botnet development. Not only can the machines be repurposed to play video games such as Doom, patient information stored in the machines' databases may be extracted, and the machines operation can be modified in such a way as to inconvenience a hospital laboratory and cause patient harm.

In addition to the morbidity and mortality caused by such integrity attacks, weaknesses in infrastructure foster a lack of trust in a connected healthcare system. Doctors, under this system, can no longer *rely* on the machines and results they use to guide decision making. With time and exposure, patients wary of insecure healthcare infrastructure may avoid seeking care. Thus, the potential clinical consequences of attacks such as those herein presented on the HL7 standard are farther-reaching than the attacks themselves. Pestilence, the tool presented in this paper, is a dead canary in a coal mine.

It is not the intention of this project to solely criticize the HL7 protocol or individual device vendors that may utilize the standard. Several newer medical devices support security controls such as encryption, but are poorly configured when deployed by healthcare delivery organizations, and do not integrate easily with existing software deployments. Updating is difficult in systems which need to operate continuously, and there is an understandable distrust of modifying systems are already working. Insecurity is complicated: it is a byproduct of culture, lack of proper education, and an overriding emphasis on interoperability without due consideration of the implications of interconnectivity. HL7 could be deployed with a secure, lower level encryption standard such as TLS, but making that change would involve every manufacturer and software provider that uses the protocol at the application level.

We propose three viable strategies to ensure increased security and integrity of data in clinical environments, which we hope will be taken into consideration by the healthcare community:

1) **Secure network deployment: network segmentation, VLANs, and firewall controls.** This is the most viable option for legacy systems and healthcare providers with budgetary and operational constraints. By restricting the attack surface of vulnerable devices to Ethernet networks inaccessible to outside influence, the potential for attack is largely mitigated. This, however, requires the intervention and trust of an experienced IT professional. When legacy devices that lack security controls exist in the network environment, isolation of these devices into network segments to minimize exposure is key.

2) **Proper configuration:** In situations where the hospital network cannot be made completely secure through use of network segmentation, the alternative is proper configuration of servers and devices that support encryption. This would mean, for example, ensuring that the interface client, such as Mirth connect, is updated to its most recent version and the communication channels are set up to use encryption.

3) **Security conscious protocols and ecosystems:** Moving forward, device manufacturers, care providers, standards organizations, and policy makers must push to incorporate newer protocols and ecosystems where strong security guarantees are built in, and actively look for these guarantees. One such example is the Fast Healthcare Interoperability Resource (FHIR), a replacement for HL7 which has greater potential for encryption [7]. Without the development of a security conscious culture, healthcare infrastructure will remain vulnerable to malefaction.

Acknowledgments:

The authors would like to graciously thank Siemens Healthcare for technical and software support of this work.Throughout the inception, creation and execution of this project Siemens has fostered a positive and supportive collaboration with our security research team.

Author's note:

The Pestilence tool is solely a proof-of-concept and will not be released to the general public. The vulnerabilities and attack methodologies discussed have been previously reported, and would-be attackers will not derive any novel assistance from this whitepaper alone.

References

(1) Caceres, J.W. & Dicorcia, M.J. Med.Sci.Educ. (2018) 28:247. https://doi.org/10.1007/s40670-017-0495-0
(2) Dolin, R., Alschuler, L. et al. J.Am.Med.Info.Assoc. (2001) 8:6:552-569 https://doi.org/10.1136/jamia.2001.0080552
(3) Spronk, R. The early history of health Level 7. Ringholm whitepaper (2014) https://bit.ly/1kkkmMg
(4) HL7 International. (2007, August 31). Implementation FAQ: Encryption and Security. (2017) https://bit.ly/2M4dUhx
(5) Haselhorst, D. HL7: Attacking and Defending the Achilles Heel of Healthcare (2017) The SANS Institute https://bit.ly/2K6uLhR
(6) Abel, R. Vanderbilt University researcher claims breaches linked to patient deaths. SC Media US (2018) https://bit.ly/2GgflGB
(7) FHIR Infrastructure Work Group, FHIR Security (2015) https://bit.ly/2LAQtAj