

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

DATAVANT, INC.

Petitioner

v.

VIGILYTICS LLC

Patent Owner

IPR2024-00311

Patent No. 10,109,375

Issue Date: Oct. 23, 2018

Title: De-Identifying Medical History Information for Medical Underwriting

**DECLARATION OF PAUL C. CLARK, DSC. IN SUPPORT OF PETITION
FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 10,109,375**

TABLE OF CONTENTS

I.	Introduction.....	3
II.	Qualifications.....	3
III.	Materials Considered	7
IV.	Legal Testimony	8
	A. Relevant Time Period for Unpatentability Analysis	8
	B. Unpatentability by Obviousness.....	8
V.	Person of Ordinary Skill in the Art.....	11
VI.	Claim Construction.....	12
VII.	Summary of Opinions.....	14
VIII.	The '375 Patent.....	15
	A. Background	15
IX.	The Grounds of Unpatentability	18
	A. Prior Art References.....	18
	B. Ground 1: Obvious (§ 103) over Evenhaim in view of Murphy.....	36
	C. Ground 2: Obvious (§ 103) over Evenhaim in view of Settimi and Murphy.....	66
	D. Ground 3: Obvious (§ 103) over Dick in view of Landi and Murphy.....	84
X.	Conclusion	117

EXHIBIT LIST

Exhibit Number	Description
1001	U.S. Patent No. 9, 965,375 (the “’375 Patent”)
1002	File History of U.S. Patent No. 9, 965,375
1003	<RESERVED>
1004	Curriculum Vitae Of Dr. Paul Clark
1005	U.S. Pat. No. 7,823,207 (“Evenhaim”)
1006	U.S. Pat. Appl. Publ. No. 2007/0294111 (“Settimi”)
1007	U.S. Pat. Appl. Publ. No. US 2002/0111833 (“Dick”)
1008	U.S. Pat. No. 7,519,591 (“Landi”)
1009	Shawn N. Murphy and Henry C. Chueh, <i>A Security Architecture for Query Tools used to Access Large Biomedical Databases</i> , 2002 AMIA Annual Symposium Proceedings, 552 (2002) (“Murphy”)
1010	<i>Milliman, Inc. v. Gradient A.I. Corp.</i> , No. 1:21-Cv-10865-NMG (D. Mass.), Dkt. 108.
1011	<i>Credit Card Security: The Best Way to Secure Data is Not to Store the Data</i> (Hospitality Upgrade, Summer 2008, www.hospitalityupgrade.com)
1012	Jennifer G. Steiner, B. Clifford Neuman, and Jeffrey I. Schiller, <i>Kerberos: An Authentication Service for Open Network Systems</i> , Winter 1988 Usenix Conference (1988).
1013	Declaration of Sylvia Hall-Ellis, Ph.D.

I, Paul Clark, state the following:

I. INTRODUCTION

1. I have been retained by Venable LLP, counsel for Datavant, Inc. (“Datavant”) to provide my opinions concerning the validity of claims 1-20 of U.S. Patent No. 10,109,375 (“the ’375 Patent”).

2. I am being compensated for my time at my normal consulting rate of \$590.00/hour. This compensation does not depend on the nature or presentation of my testimony, or on the outcome of any proceeding.

II. QUALIFICATIONS

3. I am qualified to render an expert opinion in this matter by virtue of my professional experience and education, which are summarized below.

4. In 1986, I received a Bachelor of Science degree in Mathematics from the University of California, Irvine. In 1988, I received a Master of Science degree in Electrical Engineering and Computer Science from the University of Southern California. In 1994, I received a Doctor of Science degree in Computer Science from George Washington University.

5. From 1985 to 1989, I worked as a Systems Engineer at Ultrasystems Defense and Space. As more fully set forth in my curriculum vitae attached hereto as Exhibit A, at Ultrasystems I designed and implemented large-scale simulation and network-based systems for the United States Department of Defense (DOD). A

high-speed database server I designed and implemented was used for real- time intelligence collection by the National Security Agency (NSA).

6. From 1989 to September 1990, as more fully set forth in my curriculum vitae, I worked as a Technical Lead at GTE Government Systems. While at GTE, I designed and implemented network and load generators for OS/2 LAN Manager to measure network-performance load metrics for the Central Intelligence Agency (CIA). I also developed X Windows interfaces for a largescale event-driven network system for the NSA.

7. From 1990 to 1995, as more fully set forth in my curriculum vitae, I worked as a Senior Security Engineer at Trusted Information Systems. While at Trusted Information Systems I designed and implemented high-assurance security systems, including cryptographic systems and applications for the NSA and the Defense Advanced Research Projects Agency (DARPA). My work at Trusted Information Systems involved cryptography, operating systems, intrusion detection systems, and other security technology developments and implementations including public and secret key encryption to provide security services for electronic transmissions. I am also the inventor of the Boot Integrity Token System (BITS) which provides hardware, enforced authentication within the boot sequence and guarantees operating system integrity using smart tokens. I am the named

inventor of five security related patents, U.S. Patent Nos. 5,448,045; 5,892,902; 8,695,066, 9,391,957 and 10,129,214.

8. From 1995 to 1999, as more fully set forth in my curriculum vitae, I worked as Chief Scientist at DynCorp Network Solutions, where I served as senior internal security consultant for a variety of projects. For example, I was architect and Technical Director of the IRS Secure Submission and Retrieval System that allowed the digitally signed and encrypted submission of tax data over the Internet. The successful deployment of this system resulted in three Al Gore Hammer Awards. I also created a suite of security products for providing secure wide area user access to servers that was marketed and sold to the DOD and other parts of the federal government.

9. From 1999 to the present, I have served as President and Chief Technology Officer of SecureMethods, Inc. and Paul C. Clark LLC. SecureMethods specializes in the design, implementation, and deployment of advanced secure network applications for commercial and government clients, including the DOD. SecureMethods provides a comprehensive scalable, Commercial-Off-The-Shelf (COTS) secure architecture, implemented through the use of the SM Gateway. The SM Gateway is a next-generation security appliance developed by SecureMethods that is available on UNIX-based platforms using commercial, government, and Type I cryptography, implemented in both hardware

and software. In my capacity as President and Chief Technology Officer of SecureMethods, I have technical and operational oversight of all projects and corporate technical operations. I provide guidance to senior technical personnel for design, implementation, and troubleshooting for a wide range of systems both internal and external. My work includes network systems and security, cryptographic applications, certification, key management, authentication, and integrity strategies for network applications. I also provide a wide range of high-end technical and legal consulting services. My firm specializes in complex software and hardware systems for commercial and DoD clients.

10. From 2016 to present, I have also served as CEO of ClearGuide Medical Inc. Clearguide produces and sells FDA cleared devices within the surgical community. As part of the clearance process ClearGuide Medical Inc. has successfully completed multiple clinical and non-clinical trials. As a leader in computer aided and AI guidance technologies, HIPAA compliance is central to ClearGuide Medical Inc.'s development and deployment activities.

11. I was also a member of the Federal Advisory Committee for Key Management Infrastructure (KMI), serving as Chairman of the Interoperability Working Group for Cryptographic Key Recovery. I have also served as an adjunct professor in the Computer Science Department at The George Washington University, where I have taught doctoral-level cryptography, network, and

computer security courses. I also appeared before a Congressional committee to provide testimony on “Advanced Technology for Border Control.”

12. I have co-authored a number of publications in the computer and security areas.

III. MATERIALS CONSIDERED

13. In connection with forming my opinions, I have considered the following materials:

Exhibit Number	Description
1001	U.S. Patent No. 9, 965,375 (the “’375 Patent”)
1002	File History of U.S. Patent No. 9, 965,375
1003	<RESERVED>
1004	Curriculum Vitae Of Dr. Paul Clark
1005	U.S. Pat. No. 7,823,207 (“Evenhaim”)
1006	U.S. Pat. Appl. Publ. No. 2007/0294111 (“Settimi”)
1007	U.S. Pat. Appl. Publ. No. US 2002/0111833 (“Dick”)
1008	U.S. Pat. No. 7,519,591 (“Landi”)
1009	Shawn N. Murphy and Henry C. Chueh, <i>A Security Architecture for Query Tools used to Access Large Biomedical Databases</i> , 2002 AMIA Annual Symposium Proceedings, 552 (2002) (“Murphy”)
1010	<i>Milliman, Inc. v. Gradient A.I. Corp.</i> , No. 1:21-Cv-10865-NMG (D. Mass.), Dkt. 108.
1011	<i>Credit Card Security: The Best Way to Secure Data is Not to Store the Data</i> (Hospitality Upgrade, Summer 2008, www.hospitalityupgrade.com)

1012	Jennifer G. Steiner, B. Clifford Neuman, and Jeffrey I. Schiller, <i>Kerberos: An Authentication Service for Open Network Systems</i> , Winter 1988 Usenix Conference (1988).
1013	Declaration of Sylvia Hall-Ellis, Ph.D.

IV. LEGAL TESTIMONY

14. In expressing my opinions and considering the subject matter of the claims of the '375 Patent, I am relying upon certain basic legal principles that have been provided and explained to me.

A. Relevant Time Period for Unpatentability Analysis

15. I have been asked to assume the priority date for the claims is July 1, 2009. Therefore, for the purposes of this declaration, I have analyzed obviousness from the perspective of a person of ordinary skill in the art ("POSA") as of that date.

B. Unpatentability by Obviousness

16. It is my understanding that a patent claim is unpatentable if the claimed invention as a whole would have been obvious to a POSA at the time of the invention, in view of the prior art in the field and analogous fields. I understand that a patent claim can be found unpatentable as obvious where the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the claimed subject matter was made to a POSA. I understand that an obviousness analysis involves a

consideration of (1) the scope and content of the prior art; (2) the differences between the claimed invention and the prior art; (3) the level of ordinary skill in the pertinent field; and (4) secondary considerations of non-obviousness.

17. I understand that to prove a claimed invention is unpatentable for obviousness, it may be useful to understand why a POSA would have been motivated to combine prior art references.

18. Obviousness may be shown by demonstrating that it would have been obvious to modify what is taught in a single piece of prior art, or to combine the teachings of more than one prior art reference, to create the patented invention. In determining whether a piece of prior art could have been modified or combined with other prior art or with other information within the knowledge of a POSA, the following are examples of approaches and rationales that may be considered:

- a. Combining prior art elements according to known methods to yield predictable results;
- b. Simple substitution of one known element for another to obtain predictable results;
- c. Use of a known technique to improve similar devices (methods, or products) in the same way;
- d. Applying a known technique to a known device (method, or product) ready for improvement to yield predictable results;

- e. Applying a technique or approach that would have been “obvious to try” (choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success);
- f. Known work in one field of endeavor may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations would have been predictable to one of ordinary skill in the art; or
- g. Some teaching, suggestion, or motivation in the prior art that would have led one of ordinary skill to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention. I also understand that this suggestion or motivation may come from such sources as explicit statements in the prior art, or from the knowledge or common sense of one of ordinary skill in the art.

19. I understand that an invention that might be considered an obvious variation or modification of the prior art may be considered non-obvious if one or more prior art references discourages or leads away from the line of inquiry disclosed in the reference(s). For purposes of assessing obviousness, a reference does not “teach away” simply because the reference suggests that another embodiment is better or preferred. My understanding of the doctrine of teaching away requires an indication that the combination should not be attempted (e.g.,

because it would not work or explicit statement saying the combination should not be made).

V. PERSON OF ORDINARY SKILL IN THE ART

20. I understand that an assessment of claims of the '375 patent should be undertaken from the perspective of a person of ordinary skill in the art as of the earliest claimed priority date, which I understand is July 1, 2009.

21. I have also been advised that to determine the appropriate level of a person having ordinary skill in the art, the following factors may be considered: (1) the types of problems encountered by those working in the field and prior art solutions thereto; (2) the sophistication of the technology in question, and the rapidity with which innovations occur in the field; (3) the educational level of active workers in the field; and (4) the educational level of the inventor.

22. A POSA in the field of the '375 Patent would have been a person with a Bachelor of Science degree in computer science, or a related field, and at least two years of experience in the field of computer science that includes data privacy, computer networking or data records management. Superior education could compensate for less in work experience, and vice-versa.

23. My opinions regarding the level of ordinary skill in the art are based on, among other things, my more than 30 years of experience in computer and data systems and security, my understanding of the basic qualifications that would be

relevant to an engineer or scientist tasked with investigating methods and systems in the relevant area of de-identified databases, and my familiarity with the backgrounds of colleagues and employees, both past and present.

24. By July 2009, I had Master of Science degree in Electrical Engineering and Computer Science and a Doctor of Science degree in Computer Science. I also had over twenty years of experience working in the field of computer security and database implementations, including on data privacy and computer networking. As such, I was qualified at least as a POSA, and I have had a sufficient level of knowledge, experience, and education to provide an expert opinion in the field of the '375 Patent, particularly on how a POSA would have viewed the prior art.

25. My opinions below are based on the perspective of a POSA at the time of the assumed date. I have worked on projects in the relevant field with a number of professionals who would have qualified as a POSA.

VI. CLAIM CONSTRUCTION

26. I understand that the claims must be given the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention. I applied the terms' ordinary and customary meaning in accordance with my understanding of how a POSA would understand them in July of 2009.

27. For instance, the term “healthcare information/data” is used throughout the Challenged Claims. Based on my reading of the ’375 Patent I do not believe that a POSA would limit “healthcare information” to insurance risk information. Rather, a POSA would understand that “healthcare information” to have its plain and ordinary meaning of “medical information associated with the individuals of the group.”

28. For example, claims 6 and 19 says that “healthcare data represents medical related information associated with the individuals of the group.” The specification also explains that “healthcare data” includes data “about patients, including medical and prescription drug history, and administrative healthcare claims data.” Ex. 1001 1:25-28. This construction is consistent with how a POSA would interpret “healthcare information/data.”

29. I understand that the patent owner has taken the following claim construction positions in a litigation pending in Massachusetts, *Milliman, Inc., et al. v. Gradient A.I. Corp., et al.*, Civ. Action No. 21-10865-NMG, Claim Construction Order, Docket No. 134 (Ex. 1008), which I am adopting for the purpose of my analysis:

- Token – A identifier associated with an individual that does not reveal the identity of the individual. *Id.* at 6.

- Server – A computer, with a processor and memory, on a network that provides information in response to requests from another computer.
Id. at 7-8.
- Similarly encrypted – Generated such that tokens generated from corresponding information can be compared to one another to detect a match. *Id.* at 11
- Healthcare information that characterizes a group – Plain and ordinary meaning that includes information about history of physician visits, medical conditions, procedures performed, prescriptions or disabilities. *Id.* at 13.
- Further, a POSA would understand the foregoing construction of “token” to encompass encrypted identifying information, which is stated by the ’375 Patent. Ex. 1001, 1:54-56 (“Producing such a request token ... may include encrypting respective data that identifies an individual.”).

VII. SUMMARY OF OPINIONS

30. Based on the materials that I have reviewed, the analysis I have performed, and the positions I have summarized within this report, I have formed at least the following opinions listed below.

31. Claims 1-22 of the '375 Patent are obvious over Evenhaim in view of Murphy.

32. Claims 1-22 of the '375 Patent are obvious over Evenhaim in view of Settimi and Murphy.

33. Claims 1-22 of the '375 Patent are obvious over Dick in view of Landi and Murphy.

VIII. THE '375 PATENT

A. Background

34. The '375 Patent relates to a computer-implemented system of producing information characterizing a group of individuals from a set of private data representing characteristics of individuals. De-identification is done in order to comply with the Health Insurance Portability and Accountability Act ("HIPAA"). Ex. 1001, 3:55-4:3. HIPAA requires health care providers to obtain explicit patient permission to release health records with personal identifying information to a third party. *Id.* However, obtaining such information is not always practical, especially when inquiring about the health status of a group of people. Ex. 1001, 4:4-24. For example, health insurance companies need to get the healthcare information for a group of employees in a company to evaluate the risk of insuring that group and provide an accurate price for the insurance, but getting explicit authorization from every employee when pricing health insurance with

multiple companies is not practical. *Id.* HIPAA recognizes this concern, and provides an alternative option for obtaining healthcare information by allowing for healthcare information to be sent to third parties without permission but only if the information is sufficiently de-identified before it is sent so that the identity of the individual cannot be obtained from the records. 45 C.F.R. § 164.514(b)(1).

35. HIPAA provides for two types of de-identification that allows for healthcare information to be sent to third parties without explicit authorization from the patients. The first is known as the safe harbor provision. It sets forth 18 types of information that must be removed from the patient records to send the records. 45 C.F.R. § 164.514(b)(2). Some of that barred information, like name, address, social security number, would lead to direct identification of the patient, however, some of the information barred by the safe harbor provision is less direct and might potentially be useful to health outcomes, like the birthdate of the patient. Therefore, HIPAA provides an alternative option that potentially allows more information to be included in the de-identified records sent without explicit permission. This option is referred to as statistical de-identification, and requires a determination that the information provided cannot be reverse engineered to match patient records with individuals. 45 C.F.R. § 164.514(b) (“a person with appropriate knowledge and experience with generally accepted statistical and scientific principles and methods for rendering information not individually

identifiable”). So, for example, an engineer might determine that it is acceptable to allow the age of the patient, and the geographic area where the patient lives to be included in the transmitted healthcare information, so long as there are a sufficient number of patient records so that no single patient could be matched back to each record even if a group of patients might potentially be identified from the healthcare data. The purpose of the ’375 patent is to provide a method for complying with HIPAA using a token-based system for de-identifying patient records. The limitations in the ’375 claims relating requiring a minimum number of patients are meant to comply with the statistical de-identification requirements of HIPAA.

36. Using a token system to de-identify patient records then query those records is a common technique used for storing and querying sensitive healthcare records in order to comply with Privacy laws, including HIPAA, that was well-known by the priority date of the ’375 patent. A common way to implement tokenization is to have a shared public encryption key that is made known to both the source of the healthcare information as well as the entities that want to access that information. The public encryption key is used to encrypt or tokenize one or more fields of personally identifying information that is in the record in order to create a token that is unique for each patient. The sources of patient healthcare information, such as a hospital, will typically create a de-identified version of the

patient records by applying the public encryption key and some type of encryption algorithm that cannot be reversed (or in some cases cannot be reversed without a private key) to create the unique token, insert that token to the patient records for that particular patient, and remove all personally identifying information from the record. The resulting de-identified records can then be queried using the same unique token and any matching records transmitted to a third party without violating HIPAA. In my experience, the most common technique used to query such de-identified databases has been, and was at the time of the '375 patent, to produce similarly encrypted tokens for a query based on the same personally identifying information, then transmit those tokens as the query to the server holding the de-identified records in order to match the produced tokens against corresponding tokens in the records. Producing such similarly encrypted tokens requires that the query start with personally identifying information such as name, social security number, or address.

IX. THE GROUNDS OF UNPATENTABILITY

A. Prior Art References

i. Evenhaim

37. Evenhaim is entitled “Privacy preserving data-mining protocol.” I have been informed that Evenhaim, an issued U.S. patent, was published

December 6, 2007. I have further been informed that Evenhaim is prior art under 35 U.S.C. § 102(a) or (b) (pre-AIA).

38. As discussed above, under HIPAA a party in possession of records that include the identity of the patient cannot disclose those records to a third-party without consent of the individual (45 C.F.R. § 164.502(a), 45 C.F.R. § 164.522). Evenhaim describes a protocol that would allow disclosure of privacy-sensitive information of an individual to a third party without obtaining explicit written authorization from each individual, and prevent disclosure of the identification of the individual, thereby meeting HIPPA requirements. Ex. 1005, 5:4-7.

39. Evenhaim's protocol operates between the secure aggregator data processor and at least one source-entity data processor which are interconnected via electronic data communications topology. *Id.*, 6:57-63. The source-entity data processor holds (and queries) records. *Id.*, 7:32-42. A POSA would understand that a server could provide the "electronic data communications topology" for the data aggregator and source-entity data processor. In Evenhaim the aggregator data processor first accepts a query from a user interface that includes predetermined parameters and attributes, thereby forming a parameter list. *Id.*, 6:64-67. This parameter list may include identity-disclosing specifics. *Id.*, 8:4-9. Thus, the protocol in Evenhaim describes requesting patient healthcare information from a database that includes information that identifies the patient. An encrypted token

is produced based on the request. *Id.*, Fig. 4, 12:4-13:20. I note that there are multiple possible embodiments described in Evenhaim. There is an embodiment that does not require tokens, but Evenhaim also describes using a token matching system to allow for querying de-identified patient records as shown in Fig. 4 and described in the specification. *Id.*

40. When applying tokens, Evenhaim describes creating de-identified patient records where “the person identity information, for the matched population, is replaced with a unique encryption key, such as a one way hash” *Id.* at 12:36-42. A POSA would understand that this unique key is a token. The de-identified records are then available for querying by the system shown in Fig. 4. Using this token matching system, “preserves the ability to know that two records belong to the same identity, but preserves the privacy of the identity of that person.” *Id.*, 12:42-45. Now with respect to the system shown in Fig. 4 for querying healthcare information, the specification describes that encryption can take place on the encryption server (the data aggregator in Fig. 4) as well as on the data sources. “[T]here are numerous locations where an encryption of individual identification . . . could be facilitated. By way of non-limiting example, on the ‘aggregator’ processor side, the encryption . . . could be enabled during the sub-steps: receiving a respective file” *Id.* 12:24-34. A POSA would understand that this describes that the encryption server (aggregator) shown in Fig. 4 receives

the identifiable patient list then applies an encryption key to encrypted that patient list thereby producing encrypted tokens that are distributed as a query to the data sources. Aside from the guidance in Fig. 4 a POSA would also understand that in order to comply with HIPAA, the query sent from the aggregator to the data sources should not include personally identifying information. In other words, transmitting personally identifying information to the data source as part of the query could itself be a potential HIPAA violation. Therefore, a POSA would understand that a best practice widely used in the industry, was to transmit encrypted tokens to the data sources. The data server will then match the tokens against the de-identified patient records and will return healthcare information for the group of individuals in the query. The token matching is described in the specification. “If the data sources are careful to encrypt the identifying portions of each record using a common encryption ‘key’, then further down the data processing circuit it will be possible to link records related to a common individual without compromising the identity of that individual.” *Id.*, 12:13-16. A POSA would understand that “further down the data processing circuit” describes the subsequent use of the de-identified records for querying, as shown in the system in Fig. 4. A POSA would also understand that it was common practice to compare and match the tokens as opposed to the PII. At a minimum, a POSA would understand matching tokens is preferable to matching re-identified PII because

when using the tokens the information would not need to be de-identified for the request, re-identified for the matching, and de-identified again for the response, and therefore using tokens is more efficient and further ensures compliance with HIPAA.

41. Because the tokens cannot be reverse engineered to reveal the identity of the individual, the information collected does not identify the individual. Thus, the information complies with the privacy restrictions of HIPAA and can be provided to third parties without the individual's consent. *Id.*, 8:13-18. "For example, the instant method can ask about the status of Tom, Dick, Harry, and a further collection of individuals—and obtain results from that query—so long as the statistical nature of those results will not allow correlation of a result specific back to an individual." *Id.*, 8:18-22.

42. Practically speaking, there could be an additional level of security added to the token matching systems described in both the '375 patent and in Evenhaim. It would be common and a simple design decision to add a second level of protection. So before patient records are sent back to the encryption server (data aggregator in Evenhaim) in response to the query, the tokens in the patient records could be further de-identified with a second encryption key that is only available to the data source. This allows for records belonging to the same patient to be matched for further analysis on the encryption server, but still prevents any reverse

engineering of the identity back to the tokens, and further ensure compliance with Privacy laws such as HIPAA.

43. Evenhaim's protocol is shown in Fig. 4, reproduced below.

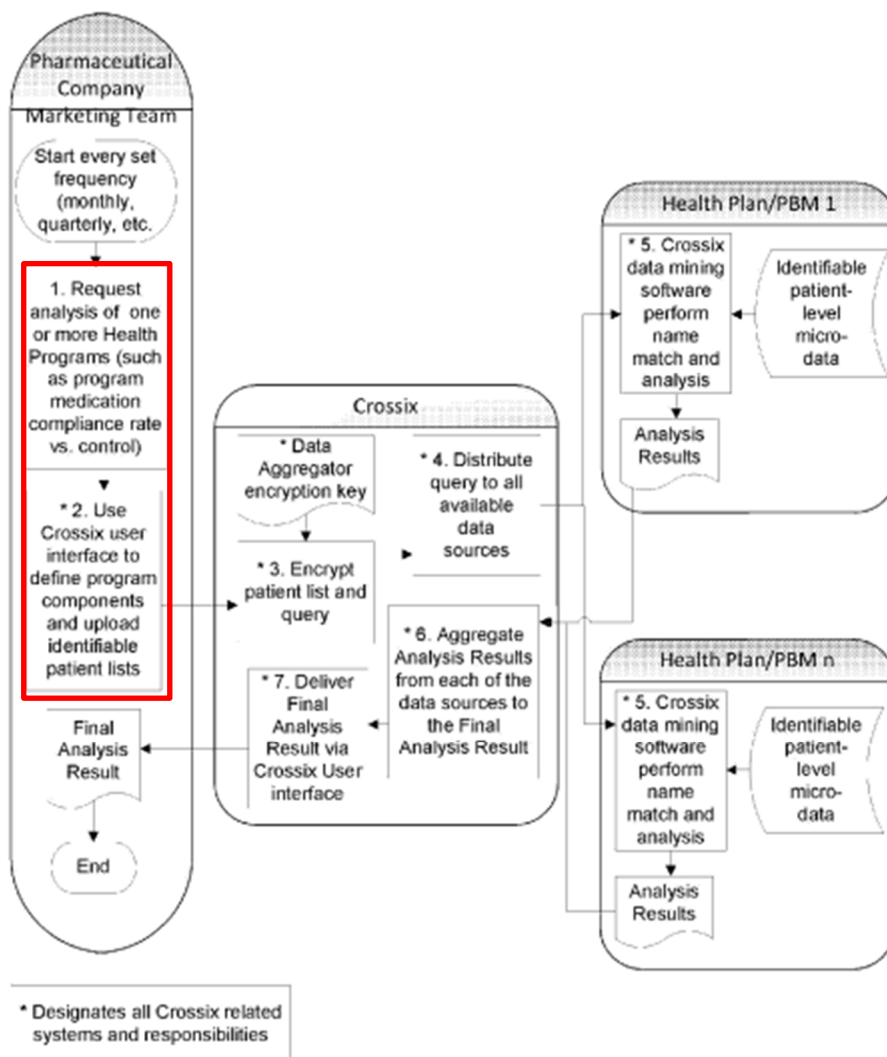


FIG. 4

44. In step 1 a user, a pharmaceutical company in this example, makes a request for an analysis of one or more Health Programs. For example, the Health Programs can request medication compliance rate versus a control. In step 2 the

user defines the program components and uploads an identifiable patient list which serves as the parameters for the query.

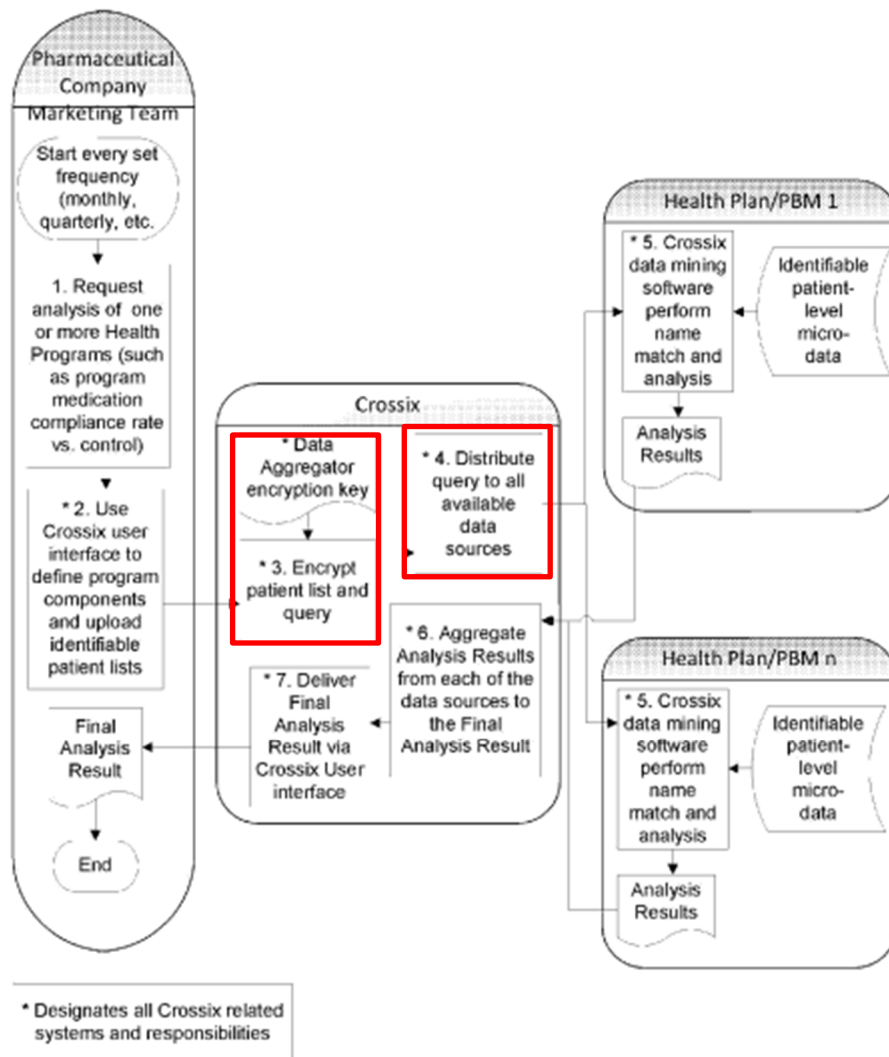


FIG. 4

45. In step 3 the data aggregator uses an encryption key to encrypt the patient list and query. In this step a POSA would understand that encryption involves creating a representation, e.g., token, for the encrypted information. The specification describes a token creation and matching system that can be

implemented at the aggregator where tokens are created by encrypting personal identifying information using a public encryption key. *Id.*, 12:4-13:20. In step 4, the encrypted query is distributed to the available data sources. This is transmitted using the topology which connects the aggregator processor and the source-entity processor. *Id.*, 6:57-63 (“Embodiments of the instant invention relate to a Privacy Preserving Data-Mining Protocol, (see FIG. 3) operating between a secure ‘aggregator’ data processor 300 and at least one secure ‘source-entity’ data processors 350, wherein the ‘aggregator’ and the ‘source-entity’ processors are interconnected via an electronic data-communications topology 399”), 7:1-2.

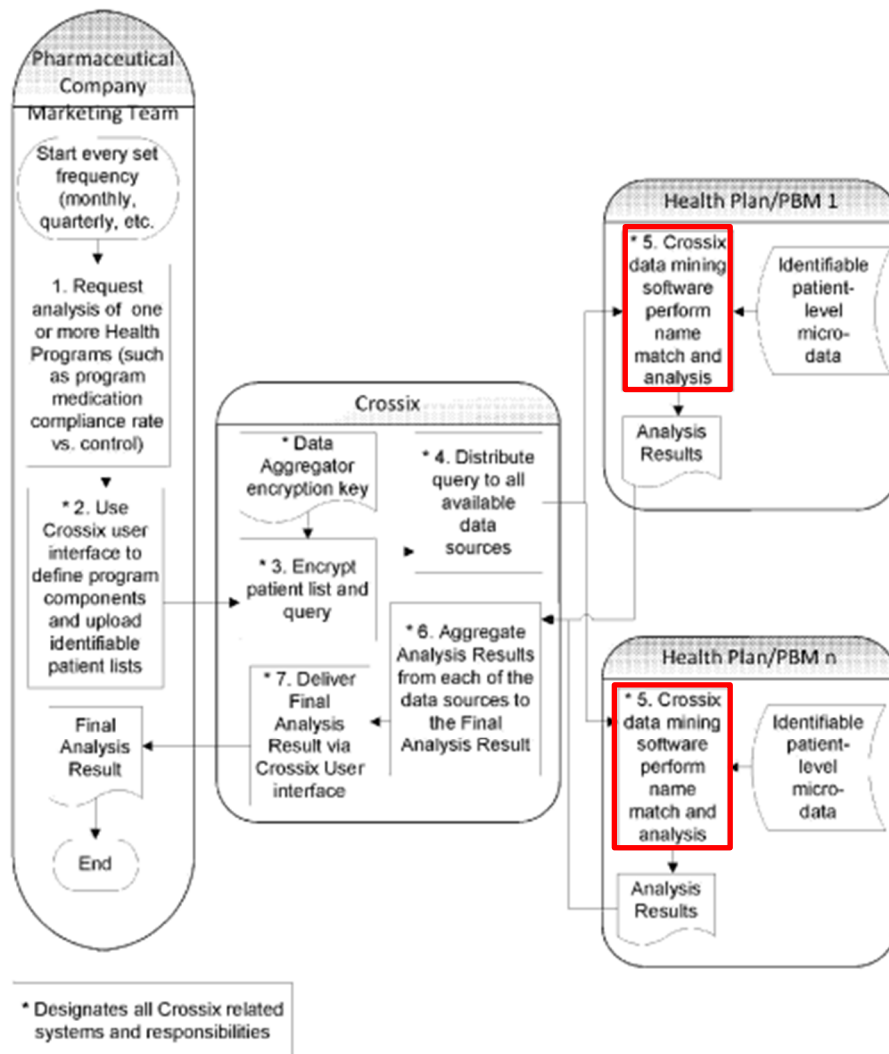


FIG. 4

46. In step 5, the program's software then performs a match analysis, e.g., matching patient records, referred to as identifiable patient level micro data in Fig. 4. Here, the patient records being queried would have been de-identified on the data source per the guidance in the specification. A POSA would understand that in the example shown in Fig. 4 a POSA would follow the specification's guidance to have the data source de-identify its patient records in order to allow those

records to be queried in compliance with HIPAA. *Id.*, 12:12-16 (“If the data sources are careful to encrypt the identifying portions of each record using a common encryption ‘key’, then further down the data processing circuit it will be possible to link records related to a common individual without compromising the identity of that individual.”). Following Fig. 4, a POSA would understand that step 5 involves matching the produced tokens sent to the data source to the corresponding tokens in the de-identified patient records which results in obtaining the patient-level micro-data that is responsive to the query. *Id.*, Fig. 4. As discussed above, Fig. 4 in view of the specification, describes or suggests that the encryption server (data aggregator) would produce tokens from the personal identifying information to query the data sources. The alternative, such as transmitting personal identifying information directly to the data sources to be used for a query, is less secure. Moreover, Evenhaim contemplates using a common encryption key to create tokens, as shown in Fig. 4, in order to match patient records “further down the processing circuit” i.e., during a subsequent query of the de-identified patient records. *Id.* 12:12-16.

47. Then, in steps 6 and 7, results from data sources’ analyses are aggregated and delivered to the user to report healthcare information back for the group of individuals that were specified in the initial query without identifying the individual, thereby complying with HIPAA. *Id.*, 7:3-4; 8:56-64.

48. As discussed further below, for ground 2, a POSA would be motivated to combine the token embodiment in Evenhaim with the central server of de-identified data in Settimi in order to simplify and enhance the querying process. Note that Evenhaim does discuss that some of the prior art systems use a central server and contrasts that scheme with how an embodiment in Evenhaim works, but that discussion is not relevant to the embodiments I discussed here, but rather is talking about the non-tokenized embodiment described in Evenhaim where the data is de-identified using only a concatenation method. *Id.* 17:28-52.

ii. Settimi

49. Settimi is entitled “Systems and Methods for Identification of Clinical Study Candidates.” I am informed that Settimi qualifies as prior art under 35 U.S.C. § 102(a) or (b) (pre-AIA) because it published on Dec. 20, 2007, which is more than one year before the effective filing date of the ’375 patent.

50. Settimi relates to using encrypted tokens to de-identifying patient data and making it available in a central repository. Ex. 1006, Figs. 1-3, [0027], [0030]. The centralized warehouse of de-identified patient data is established through a process whereby patient care providers de-identify their data when the data is updated, then transmit any newly de-identified data the centralized data warehouse. *Id.*, Fig. 2, [0027]. This central data repository allows for the de-

identified patient data to be queried for research purposes, while complying with HIPAA. *Id.*, [0027], [0030], [0035].

51. Settimi describes generating encrypted tokens called encoded patient identifiers (EPID) for each patient. *Id.*, [0041]. These tokens are inserted into the patient records in place of the personal identifying information in the records, which allows those records to be transmitted in compliance with HIPAA. *Id.*, [0039]-[0043].

52. Once in the central repository (the data warehouse), the de-identified healthcare records can be queried using the encrypted tokens, by matching the tokens in the healthcare records to tokens generated in connection with a query for healthcare data for a group of individuals. *Id.*, [0037]. These queries can be used for various types of research. *Id.*, [0043].

iii. Dick

53. I understand that Dick is a U.S. patent application that published on August 15, 2002, is therefore prior art under 35 U.S.C. § 102(a) or (b) (pre-AIA).

54. Dick describes a method for searching for medical information, including receiving a request for medical information relating to a specific subject, transmitting a query to a medical information repository (e.g., a physician's office) for that information, and receiving a response to the query. Ex. 1007, at Abstract. Further, the response to the query may have de-identified healthcare information so

that it is void of any identifying information and complies with HIPAA regulations. *Id.*, Abstract, [0013]. Dick recognizes that identifying information may be removed from the medical information at a variety of possible places, including at the site that originates the data, but the important point is that the data must be de-identified at some point before it is distributed. *Id.*, [0045]-[0046]. Dick uses de-identifying methods to allow information to be distributed for medical research. *Id.*, [0046]. In other words, Dick teaches a POSA how to use particular criteria to search for a group of individuals, and then return a de-identified data set. *Id.*, [0010]-[0013]. Dick also describes using its method in calculating the risk for healthcare insurance provides for an individual or group of individuals through de-identified data. *Id.*, [0003], [0037]. As the claims of Dick make clear, and based on what a POSA would know, it was well-known in the art to request medical information for a group of individuals, search a medical information repository, generate a response based on the search, whereby the search returns de-identified healthcare information to the requester in order to comply with HIPAA. *Id.*, Claims 23-24.

55. While Dick describes that known “de-identifying methods and processes” can be used to de-identify the healthcare data involved in order to comply with HIPAA, it does not explicitly set forth any particular method. *Id.*,

[0013], [0044]-[0046]. Dick further describes that de-identification can take place in various places, including at the originating source of the information. *Id.*

iv. Landi

56. Landi is entitled “Systems and Methods for Encryption-Based De-Identification of Protected Health Information.” I am informed that Landi qualifies as prior art under 35 U.S.C. § 102(a) or (b) (pre-AIA) because it issued on April 14, 2009, which is before the effective filing date of the ’375 patent.

57. Landi was not considered by the Patent Office during prosecution of the ’375 patent.

58. Generally speaking, Landi relates to systems and methods for protecting patient privacy when health care information is shared between various healthcare entities to maintain compliance with HIPAA regulations by using a token generation and matching system to de-identify patient data. Ex. 1008, 1:14-16; 2:30-34; 3:52-57. Landi does this by generating an encrypted ID for individuals which represents one or more items of individual identifying information. *Id.*, 2:46-49.

59. Landi’s invention is further described in the Abstract:

Systems and methods are provided for protecting individual privacy (e.g., patient privacy) when individual data records (e.g., patient data records) are shared between various entities (e.g., healthcare entities). In one aspect, systems and methods are provided which implement secured key encryption for de-identifying patient data to ensure patient privacy, while allowing only the owners of the patient data or legally

empowered entities to re-identify subject patients associated with de-identified patient data records, when needed.

60. Landi describes a system for protecting patient privacy when patient data is shared between a number of different entities. The system described in Landi “comprises a plurality of patient data processing systems (20-1~20-n), a central patient data processing system (30), a trusted broker system (40), and a central surveillance system (50).” *Id.* at 4:15-19. Landi’s system is depicted in Figure 1 below:

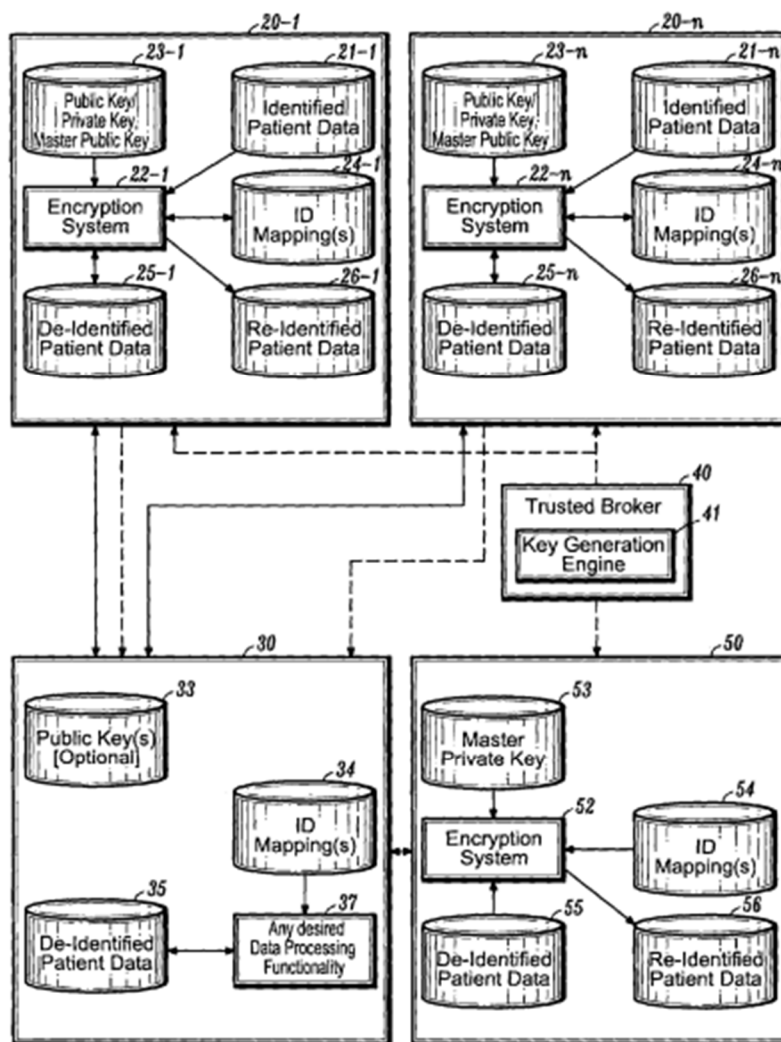


FIG. 1

61. The system described in Landi includes a plurality of patient data processing systems, a central patient data processing system, a trusted broker system, and a central surveillance system. *Id.*, 4:15-19. The patient data processing systems are different healthcare organizations that release de-identified data to the central patient data processing system where they can be processed for, e.g., research and healthcare monitoring. *Id.*, 4:19-28. This system is implemented using

an encryption scheme for protecting patient privacy wherein keys are used to encrypt and decrypt patient data. *Id.*, 4:35-48. A POSA would understand that using a key to encrypt the patient data would result in an encrypted token that is used to replace personal identifying information in the patient records. In Landi the trusted broker is a third-party who generates the encryption keys, which are securely disseminated to the appropriate entities for purposes of de-identifying and, in some specific cases, re-identifying patient data. *Id.*, at 8:5-10. In this way Landi provides both public and private keys, whereby the public keys can be used to generate tokens from personal patient information and match that token to de-identified patient records which also have a similarly encrypted token. The private key may be used to re-identify the patient from the token in certain circumstances and may be held by the trusted third party. Landi also describes an embodiment where the tokens are created using a one-way hash, and a POSA would understand that using this type of encryption would likely mean that the personal identifying information is not recoverable from the token. In some embodiments, Landi's central surveillance system is an entity that is authorized to re-identify subject patients that are associated with de-identified patient data maintained and processed by the central patient data processing system. *Id.*, at 4:28-34.

62. Each patient data processing system includes persistently stored electronic patient data records containing identified patient data, an encryption system, and optional securely stored encryption keys. *Id.*, 4:49-54.

63. The Encrypted ID of a patient is an encrypted version of patient identifying information such that it can be used to uniquely identify the patient. *Id.*, 5:13-15. The Encrypted IDs can be included in de-identified patient records, which are released to third-party entities, to correlate records without compromising patient privacy. *Id.*, 5:41-44.

64. Landi also provides a real-world example for a possible implementation: monitoring disease outbreaks. *Id.*, 8:64-67. In this example, the central data processing system is operated by a government agency that collects de-identified patient records from clinical sites. *Id.*, 8:67-9:6. A POSA would understand that this is done by setting certain parameters and searching the clinical sites for matches. The government can monitor an outbreak by collecting and analyzing de-identified patient records while maintaining patient privacy. *Id.*, 9:6-10. The government analyzes the de-identified patient data to determine if an outbreak has occurred. *Id.*, 9:10-16. Another example Landi provides is monitoring patients in a clinical trials or research studies. *Id.*, 9:43-60.

v. Murphy

65. Murphy is an article titled “A Security Architecture for Query Tools used to Access Large Biomedical Databases” authored by Shawn N. Murphy, MD, Ph.D., and Henry C. Chueh, MD, M.S. The article was published in November 2002, in the AMIA Annual Symposium Proceedings. I have been informed that Murphy is prior art under 35 U.S.C. § 102(a) or (b) (pre-AIA).

66. Murphy was the result of a study which sought to prevent the identification of individuals through the use of a search on a medical database, even when a user was trying to identify an individual. Ex. 1009, p. 553. In other words, the study sought to prevent a user from using specific search parameters and filters to associate medical records with an individual. *Id.*, p. 555. One way Murphy did this was by ensuring that when a search is performed the search result cannot identify only one individual. *Id.*, p. 555. To do so Murphy describes an algorithm which blocks any queries that return fewer than 3 patients in so as to prevent the user from finding the patient identity *Id.*, p. 555. A POSA would understand that, for example, allowing a search for one patient would not mask the identity of the patient, even if the results were de-identified. A POSA would therefore understand that allowing a search for one patient would not be in accordance with HIPAA’s concern with preventing the identification of the patient associated with any medial record.

B. Ground 1: Obvious (§ 103) over Evenhaim in view of Murphy

67. In my opinion, each element of the '375 patent's claims 1-22 would be obvious to a POSA over Evenhaim in view of Murphy. Support for my opinions is detailed further below.

68. A POSA would look to combine Evenhaim with Murphy because Murphy provides additional information about protecting the identity of individuals by preventing their identity from being revealed when a parameter-based search is performed on a medical database in compliance with HIPAA. This aligns with one of the goals of Evenhaim. More specifically, Murphy more fully explores the security issues related trying to identifying a patient based the results of query that receives de-identified medical information, something a POSA implementing the invention in Evenhaim would be motivated to use in coming up with a secure system that complied with HIPAA and other privacy regulations.

69. Murphy describes the results of a study which sought to prevent the identification of individuals in a search on a medical database. Ex. 1009, p. 553. One way Murphy did this was by ensuring that more than one individual is returned when a search is performed. Id., p. 555. Murphy describes an algorithm which blocks queries that return fewer than 3 patients thereby preventing disclosing the patient's identity. Id., p. 555.

70. A POSA would have a reasonable expectation of success combining Evenhaim and Murphy because they are all directed to searching for sensitive

information in a database and providing that information in a way that an individual cannot be identified. Evenhaim and Murphy are performed in similar computing environments so a POSA would be able to combine the references with a reasonable expectation of success. Moreover, making a modification according to what Murphy describes, i.e., ensuring a minimum number of results is returned, is a simple design choice for a POSA to make when seeking access to the healthcare information.

i. [7.a] A system comprising: a data server device

71. Evenhaim describes a data server, referred to as a “data source” or “source entity.” Figs. 4 and 4A show the source entity records are analyzed, aggregated, and delivered to the user. The source entity is described as being a computer processor having software that allows it to execute its protocol. Ex. 1005, 17:53-63, 5:12-28. The data source processors receive requests from the encryption server (data aggregator) and respond to those requests by producing the healthcare information that characterizes the group, such as the de-identified patient records. *Id.*, Figs. 3, 4, 4A, 5; 12:13-16, 7:53-55 (“Now, the ‘aggregator’ and the ‘source-entity’ processors are interconnected via an electronic data-communications topology—such as the Internet.”).

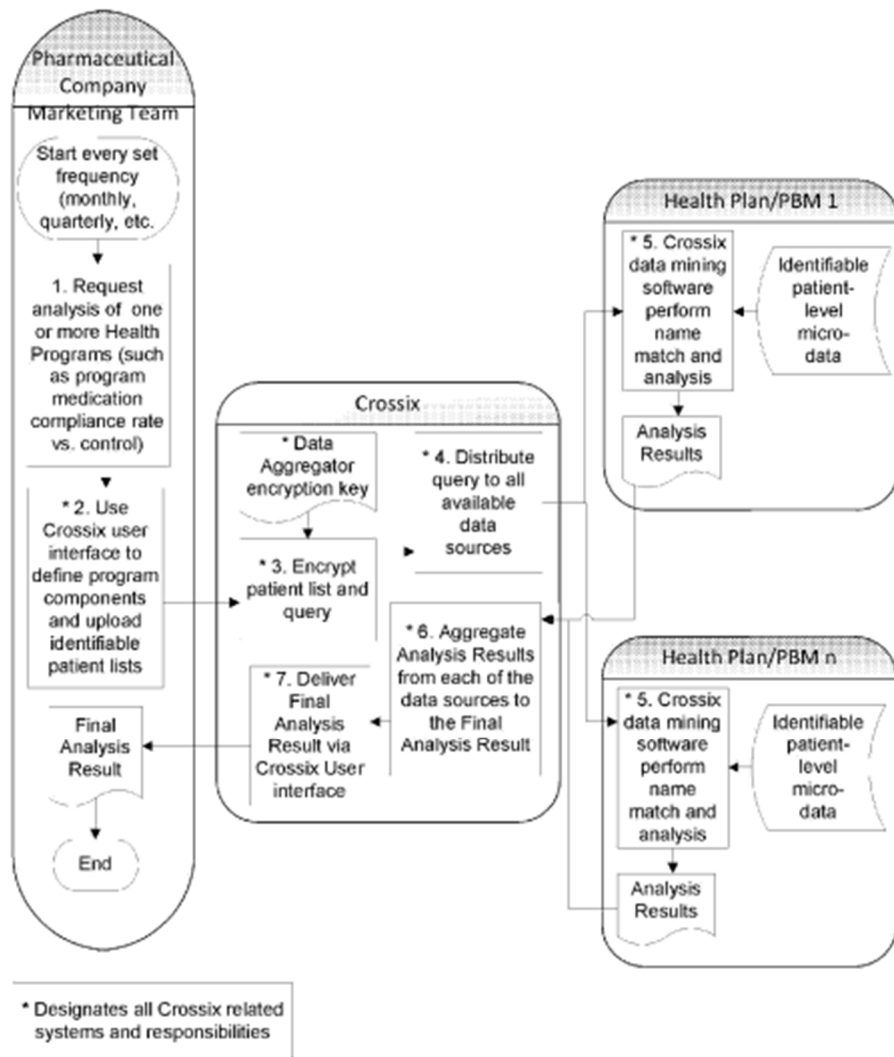


FIG. 4

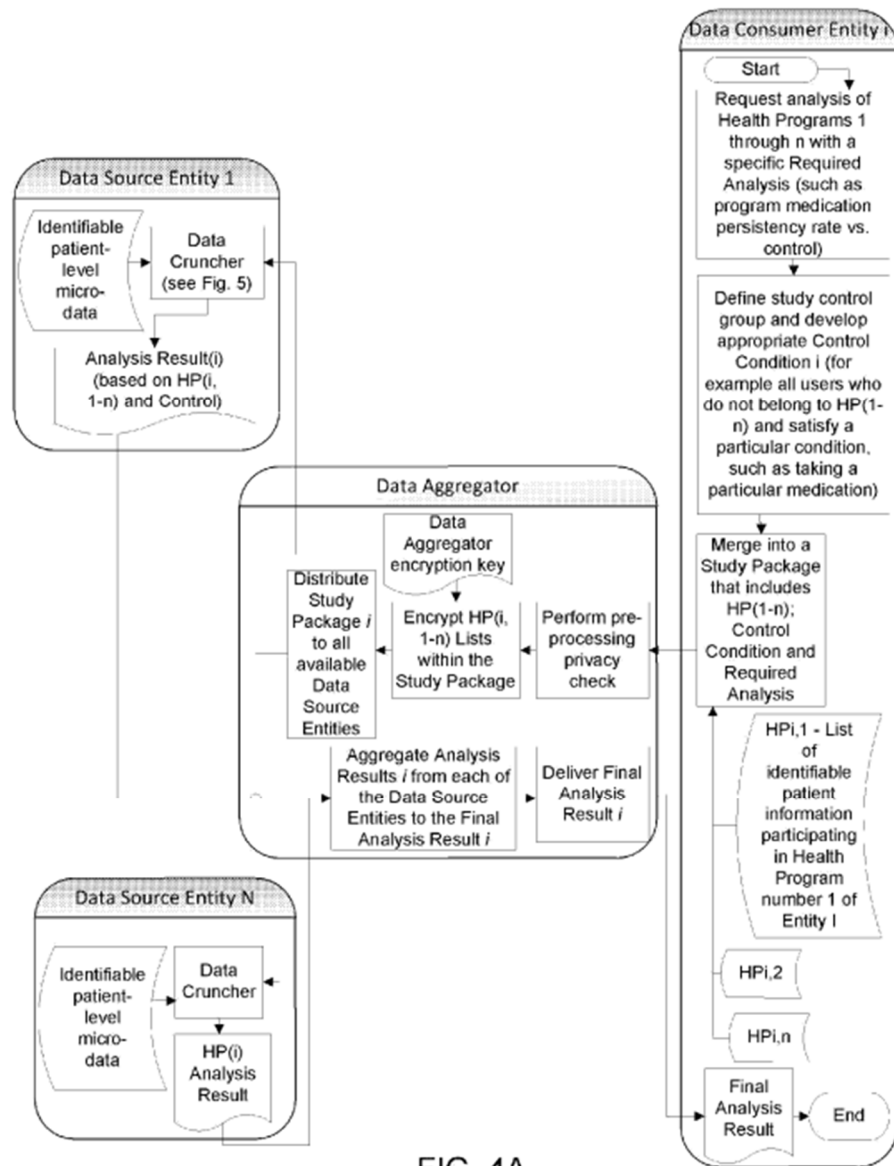


FIG. 4A

72. Data processors and data servers are known in the art as being analogous. For example, servers are data processors that respond to requests, like the data processor described in Evenhaim. A POSA therefore would have understood and known that data processors which generate and transmit data over the network suggest the use of a data server for producing a request.

73. Additionally, it would have been obvious to a POSA, that Evenhaim's data processor is the same as a data server. And it would have been obvious to a POSA that Evenhaim's data processor is used the same way as a data server.

74. The data sources can provide various types of health information that characterizes the group, including the de-identified patient records for listed patients. *Id.*, Fig. 5, 7:12-26. Other examples of health information it produces are common treating physicians for the group, tracking the health outcomes of patients in a clinical trial, and treatment data. *Id.*, 8:65-9:12; 15:3-14; 15:25-45.

ii. [7.b] for producing healthcare information that characterizes a group of a number of individuals identified in a request for the healthcare information that characterizes the group

75. As explained above, Evenhaim describes that the data source that can provide health information that characterizes a group in response to requests from an encryption server (data aggregator). Ex. 1005, Figs. 4, 4A, 5, 12:13-16, 12:42-45.

76. Evenhaim's data server provides various types of health information that characterizes the group, including the de-identified patient records for listed patients, treating physicians, health outcomes, etc. Ex. 1005, 8:65-9:12, 15:3-14, 15:25-45, 8:49-64 ("the aggregator receives some answer (or a null answer) from each source-entity"), 6:63-7:27.

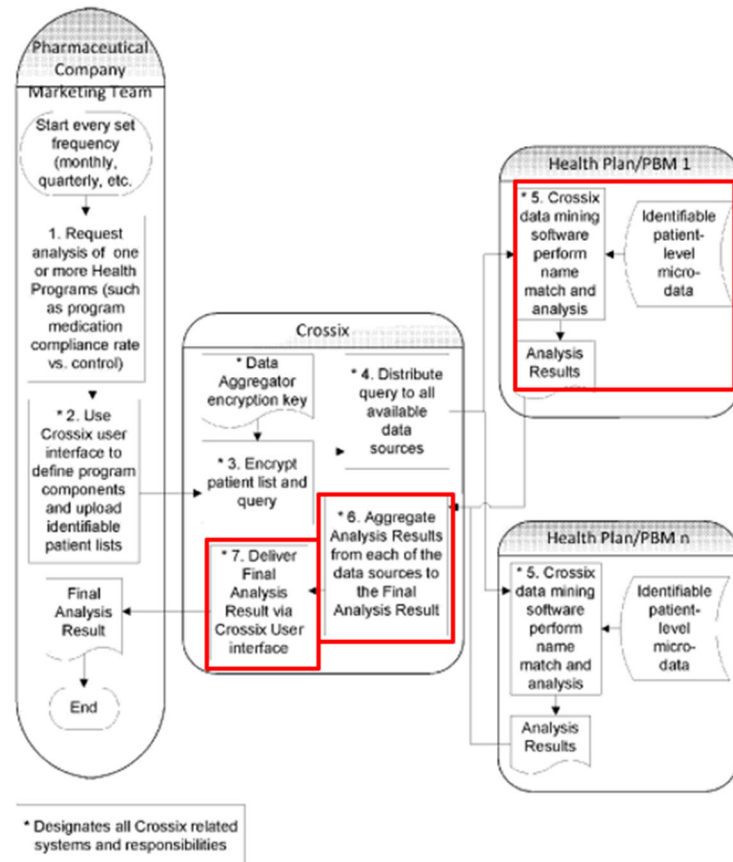


FIG. 4

77. Evenhaim explains that the data server delivers the healthcare information that characterizes the group in response to a request that is received by the data aggregator. The analysis shown in boxes 5-7 of Figure 4 is performed in response to the request for analysis shown in box 1 of Figure 4. *Id.*, Fig. 4 (showing that a user “request[s] analysis of one or more Health Programs”), 16:8-10 (“the query engine runs the analysis of their health transaction data to determine the result of the requested analysis”).

- iii. [7.c] wherein the request for healthcare information is sent to an encryption server that is configured to determine if the

number of individuals in the request is at least a defined minimum number of individuals,

78. Evenhaim describes an encryption server. For example, Evenhaim describes that its protocol operates between a secure aggregator data processor and at least one secure source-entity data processor, which are interconnected via an electronic data-communications topology which have a memory. An “electronic data-communications topology” includes the use of a server. As Evenhaim describes, the topology may include the Internet, or a virtual private circuit or the likes. Ex. 1005 at 6:57-63, 7:53-58.

79. Evenhaim describes that the aggregator and other machines are “interconnected with each other using standard data-communications facilities (e.g. Internet . . .).” *Id.*, 6:34-42. The aggregator and source-entity processors can both perform encryption. *Id.*, 12:24-34. Evenhaim also describes that its processes can be implemented on a standard computer using software packages stored on a program storage device. *Id.*, 6:34-45, 10:64-11:2, 17:53-59.

80. Evenhaim describes that in certain embodiments, a minimum number of individuals must be included in a request such that any result may not be reverse engineered to the identity of an individual. Ex. 1005, 8:13-18, 13:62-65 (“the inclusion of at least a predetermined minimum number of individuals . . . into a ‘parameter list.’”); *see also id.* 8:27-29, 16:26-32. This is required by the HIPAA Expert Determination method (an alternative to using the Safe Harbor rule as

discussed above in the background art), with which Evenhaim is intended to comply. 45 C.F.R. § 164.514(b)(1); Ex. 1005, 8:13-18 (“will present sufficient statistical distribution to protect the ‘reverse engineering’ of the result back to any individual”). One obvious way that a POSA would implement requiring a minimum number of individuals in the request is to determine at the encryption server whether there was a minimum number of individuals included in the request, as suggested by Evenhaim. *See* Ex. 1005, 13:62-65. A POSA would have understood that ensure that a parameter list or population is sufficiently large to protect privacy by determining whether a minimum number of individuals are included in the request would be a simple design choice to comply with HIPAA.

81. Since a request is initiated by a parameter list being sent to the data aggregator, it would be obvious to a POSA that the check for whether the request contains a minimum number of individuals could occur by the aggregator after the parameter list is received, or on any of the other servers in the system. Evenhaim explicitly discloses that either the aggregator or the data sources can execute steps of its protocol. Ex. 1005, 17:53-60.

iv. [7.d] wherein a unique request token is produced for each individual included in the group

82. Evenhaim describes utilizing an encryption key for creating a token, including in the encryption server, called the data aggregator. Ex. 1005, 12:4-13:20 A POSA would understand the system in Fig. 4 describes the token creation and

matching system. *Id.* A request for healthcare information is created in steps 1 and 2 using a list of patients. *Id.*, 8:18-23 (“[T]he instant method can ask about the status of Tom, Dick, Harry, and a further collection of individuals – and obtain results from that query – so long as the statistical character of those results will not allow correlation of the result[] back to an individual.”). The request is then sent to the data aggregator, in step 3. A POSA would understand that Fig. 4 shows that an encryption key can be applied to the personal identifying information in the request (such as the names of the patients or their personal identities) to create an encrypted token that can be used to query de-identified patient data. *Id.*, 8:4-22; 6:64-67. The specification describes that tokens can be created on the aggregator as well as on the data sources. *Id.*, 12:24-34. Therefore, Evenhaim suggests creating an encrypted token for each individual included in a group composed of multiple individuals identified in a request for healthcare information that characterizes the group. The token is created by encrypting the identifying information using an encryption key. *Id.*, 12:4-16, 12:38-45, 13:4-12.

83. Such tokens are computed data values used to verify an identity without disclosing potentially sensitive information regarding the identity. Such software generated tokens have been in common use since the 1980’s. *See Ex. 1012.* Software tokens were later incorporated and supported by Microsoft Windows 2000.

84. Evenhaim suggests encrypting personally identifiable information for individuals in a group and replaces it with a unique encryption key. *Id.*, 12:36-45, Fig. 4. A POSA would understand this includes a representation of data, which is a token as discussed in the '375 patent.

85. Evenhaim also describes that a user can request the healthcare information that characterizes a group of individuals. *Id.*, Figs. 4, 4A, 5, 7:65-8:4, 8:18-22, 8:30-41, 8:65-9:12, 15:3-14, 15:25-45.

86. Thus, Evenhaim describes replacing personal identity information in a population, requesting searches regarding properties of a group of individuals (requesting a query), and obtaining results from that query, so long as those results do not allow correlation of a specific result to an individual. Replacing the personal identity information data is the same as using a token in the '375 patent.

v. [7.e] the data server is configured to compare the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens,

87. Evenhaim suggests that the data server is configured to compare the request tokens to the tokens associated with healthcare data on the data server. For example, Evenhaim describes that the one-way hash or Advanced Encryption Standard preserves the privacy of the identity of an individual while also being able to know that two records (e.g., a record of the individual and a record that preserves the privacy of the individual) belong to the same individual. *Id.*

Evenhaim also describes non-token embodiments of the invention where records are de-identified by concatenating the results of the search, e.g., in 17:41-52, but that is irrelevant to the token embodiment shown in Fig. 4 and described in column 12.

88. It would also have been obvious to a POSA to implement a token creation and matching scheme in the system shown in Fig. 4. Fig. 4 shows the use of an encryption key to encrypt the patient identifying information. A POSA would understand that a known way to do that is to create encrypted tokens to match against tokens used to de-identify patient records in the data sources (such as at a hospital or doctor's office), as described in the specification. *Id.*, 12:4-13:20. As shown by Fig. 4, the encrypted tokens created for the query are then sent to the data sources (box 4) to obtain corresponding de-identified micro-level patient data back from the data sources in box 6. *Id.* Evenhaim describes that the token creation could be performed on the aggregator as well as on the data sources, and specifically calls for the encryption to occur at the data sources. *Id.*, 12:24-34.

89. A POSA would understand that each token in a set of de-identified healthcare data is associated with healthcare data for a corresponding individual in the system shown in Fig. 4 of Evenhaim. "[T]he data sources are careful to encrypt the identifying portions of each record using a common encryption 'key'" so that each individual's healthcare record is de-identified with a token (e.g., the

encrypted identifying information). Ex. 1005, 12:8-16, 12:36-45 (“[T]he person identity information, for the matched population, is replaced with a unique encryption key This . . . preserves the ability to know that the two records belong to the same identity. . .”).

90. Additionally, Evenhaim describes that the prior art also used a token matching system for de-identifying patient records and allowing those records to be queried. In Figures 1-2, medical data such as medical, hospital, and pharmacy claims data is merged into a file and then de-identified using a one-way hash or other mechanism that would have the same id for multiple records. That data can then be queried using tokens that were created using the same or similar encryption key. *Id.*, 16:46-58; 17:27-41. Further, because, as Evenhaim acknowledges, it was known to de-identify patient records prior to searching, a POSA would find it obvious that this de-identification could take place before or after receiving a request to search the records. *Id.*, 17:27-41.

91. Evenhaim refers to prior art de-identification systems, such as the one used by IMS (one of the largest providers of healthcare information), that implemented these techniques to comply with HIPAA. I was in the industry at the time, and was long familiar with the prior art systems for de-identifying data, including the one used by IMS. Any such DBMS that maintains information related to sensitive data fields requires tokenization of those fields using

cryptographic techniques. As a result, to run a query against the IMS data for a group of patients, tokens were created by a server that would run the query against a data server, in accordance with the prior art system shown in Figs. 1 and 2. *Id.* The querying tokens and the encrypted tokens would typically be created using the same public encryption key and therefore would be able to be matched, while never transmitting the personal identifying information in order to comply with HIPAA.

92. It would have been obvious to a POSA that because the identity of the individuals is unidentifiable from the tokens, the data server would compare the request tokens with the tokens associated with healthcare data in order to identify which healthcare data is associated with the requested individuals.

vi. [7.f] the produced healthcare information is absent personally identifiable information and

93. Evenhaim explains that the healthcare produced using its invention is absent personally identifiable information (is de-identified) to comply with HIPAA regulations by using the tokens, and in addition, filtering out details of the file so that portions of the answer which might allow the user to guess the identity some data attributes are eliminated. *Id.*, 8:49-64, 9:13-18, 10:1-6, 12:12-20, 12:35-44, 14:8-13.

94. The prior art system shown in Figs. 1 & 2 also shows a system for producing de-identified healthcare information. *Id.*, 16:46-57, 17:27-41.

vii. [7.g] is based on a minimum number of matched tokens,

95. As discussed above, Evenhaim describes that in certain embodiments, a minimum amount of data must be identified to produce a sufficient statistical distribution before it is produced such that any result may not be reverse engineered to the identity of an individual. Ex. 1005, 8:13-18, 13:62-65 (“the inclusion of at least a predetermined minimum number of individuals”); *see also id.* 8:27-29, 16:26-32. This is required by the HIPAA Expert Determination method (an alternative to using the Safe Harbor rule as discussed above in the background art), with which Evenhaim is intended to comply. 45 C.F.R. § 164.514(b)(1); Ex. 1005, 8:13-18 (“will present sufficient statistical distribution to protect the ‘reverse engineering’ of the result back to any individual”). One obvious way that a POSA would implement requiring a minimum number of individuals in the query results is to require a minimum number of corresponding tokens in the query. Each person identified in the query will correspond to a produced token, and it would be an obvious design choice where to enforce the requirement of a minimum number of tokens or individuals, either on the requesting side (the aggregator), which creates the produced tokens, or on the data server side, which stores the corresponding tokens.

96. Evenhaim further explains that the number of individuals in a query response must be of a number that protects privacy. Ex. 1005, 8:43-48 (“It will

simply be necessary that the number of the individuals in any population definition be large enough to dilute any definitive conclusion about an individual's personal data into the 'sea' of the group (of names') data."), 16:5-32 ("The query engine determines whether the number of participants in each program and the % participants matched are above a level that ensures consumer individual privacy (based on a statistical definition)."). Evenhaim also acknowledges that it "allow[s for] either processor to execute characterizing steps of the protocol. Ex. 1005, 17:53-60. Therefore, a POSA would consider enforcing a minimum number of matching tokens, which ensures that the minimum number of tokens are associated with patient data. This would be a simple design choice to comply with HIPAA.

97. Evenhaim in combination with Murphy renders this limitation obvious. Murphy provides more detail about statistical de-identification and subsequent querying of that data, describes this can be accomplished in compliance with HIPAA to prevent reverse engineering of the patient identity from the produced de-identified healthcare information. Ex. 1009, p. 555 ("One vulnerability is that we allow a large number of queries to be performed that return zero patients. Unfortunately, naïve users who are performing incorrect queries will often have zero patients returned. Having the table-watching agent terminate access to these users would be too sensitive a response. However, attacks around zero are more likely to succeed. Therefore, we modify our range algorithm to have

a wider obfuscation at values close to zero, and indicate the greater range by returning '<3' when the obfuscated number is less than 3"). Murphy recognized the potential for misuse when an individual searches using very specific parameters to intentionally return a small number of individuals. To combat this problem, Murphy describes an algorithm designed to prevent the user from determining the identity of the individual patients. One of the safeguards used is to block any queries that return fewer than 3 patients in order to prevent the user from finding the patient identity. It would have been obvious that Murphy's algorithm could have been combined with Evenhaim to ensure that a minimum number of tokens are identified because Murphy provides more detail on the same concern about statistical de-identification described in Evenhaim, and therefore a POSA would be motivated to combine Murphy with Evenhaim with a reasonable expectation of success.

98. Data stored on data servers such as Evenhaim's data source is susceptible to many different types of attacks. Certain attacks can be designed to return a small number of individuals in order to determine the individuals' identities, while others can be designed to return a greater number of individuals, but use targeted specific requests, in combination with known information, in order to determine the individuals' identities. Therefore, it would be obvious to a POSA to check both that a minimum number of individuals have been requested and that

a minimum number of individuals are returned to protect against each of these types of attacks in order to comply with HIPAA.

99. It would have been a simple design choice with predictable results to determine whether the number of individuals returned was at least a minimum number upon receiving the request, upon matching the tokens, or both to ensure compliance with HIPAA.

viii. [7.h] the data server is also configured to provide the produced healthcare information to report the characteristics of the group

100. Evenhaim describes that its data source processors receive requests from the encryption server (data aggregator) and respond to those requests by producing the healthcare information that characterizes the group, such as the de-identified patient records. *Id.*, Figs. 3, 4, 4A, 5; 12:13-16, 7:53-55. Evenhaim describes providing produced healthcare information to report the characteristics of the group, in the form of reporting the medical information for the group. *Id.*, Fig. 4 (*see* boxes 6 and 7), 10:6-11. The reported healthcare information can include the de-identified patient records themselves, *id.*, Fig. 5, 7:12-26, as well identifying common treating physicians for the group, tracking the health outcomes of patients in a clinical trial, as some examples described in Evenhaim. *Id.*, 8:65-9:12, 15:3-14, 15:25-45.

ix. [7.i] absent authorization from the individuals in the group.

101. Evenhaim’s system describes or suggests producing healthcare information without individuals in the group having authorized its release. As the ’375 patent recognizes, under HIPAA protected health information can only be released without permission when it is irreversibly de-identified. 45 C.F.R. § 164.502(d). A POSA would have understood that the purpose of the invention in Evenhaim is to produce de-identified healthcare data in a manner that is compliant with HIPAA, as an alternative to obtaining explicit patient permission (in which case it would not be necessary to de-identify the healthcare records). Ex. 1005, 3:17-31, 4:1-3, 14:7-13, 4:34-48, 5:46-6:10, 8:18-22, 14:7-13, 14:59-67.

- x. **[8] The system of claim 7, wherein the provided healthcare information is in conformity with the Health Insurance Portability and Accountability Act (HIPAA).**

102. The Evenhaim disclosures cited for 7.i also show that claim 8 is obvious.

- xi. **[9] The system of claim 7, wherein the request for healthcare information that characterizes a group of individuals includes personally identifiable information associated with the individuals of the group.**

103. Evenhaim describes or suggests that the request for healthcare information that characterizes a group can include personally identifiable information associated with members of the group (e.g., their names). For example, Evenhaim describes that the request may include a “parameter list,” which may include names, dates of birth, etc. Ex. 1005, 13:62-14:13, 8:39-43

(“[T]he parameter list may even include specific individuals by name or ID or the like[].”), 8:18-22 (“For example, the instant method can ask about the status of Tom, Dick, Harry”), 12:38-45, 12:8-16.

xii. [10] The system of claim 7, in which the data that identifies each individual is encrypted.

104. Evenhaim describes producing healthcare information where the personally identifying information has been encrypted. Ex. 1005, Fig. 4, 4A, 5, 12:4-20 (“[E]ncrypt the identifying portions of each record using a common encryption ‘key’, then further down the data processing circuit it will be possible to link records related to a common individual without compromising the identity of that individual.”). A POSA would understand that one of the ways Evenhaim describes providing de-identified data is by substituting encrypted tokens into the patient records being produced in response to the query in place of the personally identifying information. *Id.*, 12:4-13:20. The encrypted tokens are created by encrypting personally identifying information. *Id.* As Evenhaim explains, encryption of individual identifying information can be facilitated at many different locations, including at the aggregator and source-entity side. *Id.*, 12:24-34. Thus, a POSA would understand that producing the requested file includes encrypting data within it.

105. Additionally, the prior art system shown in Figs. 1 & 2 also show producing de-identified healthcare information involves encrypting the personal identifying information. *Id.*, Figs. 1-2, 16:46-57, 17:27-41.

106. Further, to the extent it is argued that Evenhaim does not explicitly teach this element, it would have been obvious to a POSA to encrypt data that identifies each individual to further protect that individual's personal identifying information from inadvertent disclosure in view of Evenhaim and HIPAA regulations.

xiii. [11] The system of claim 7, in which the tokens associated with the healthcare data and the request tokens are similarly encrypted.

107. Evenhaim describes or suggests that the produced tokens and tokens in a set of de-identified data are both similarly encrypted such that they can be matched. Ex. 1005, 12:39-44 (“compatible encryption”), 12:4-16 (“common encryption ‘key’”). Common encryption is used, at least, in order to ensure that records for the same patient can be matched. *Id.*, 12:12-16 (“If the data sources are careful to encrypt the identifying portions of each record using a common encryption ‘key’, then further down the data processing circuit it will be possible to link records related to a common individual without compromising the identify of that individual.”), 12:25-30 (“[T]here are numerous locations where an encryption of individual identification and/or merger of data for individuals could be

facilitated. By way of non-limiting example, on the ‘aggregator’ processor side . . .
.”).

108. Evenhaim also shows a prior art system that a POSA would understand also has produced tokens similarly encrypted to matching tokens in a set of de-identified patient records. *Id.*, Figs. 1 & 2, 16:46-57, 17:27-41.

- xiv. **[12] The system of claim 7, in which the data server provides the requested healthcare information if the group includes at least the minimum number of individuals.**

109. The Evenhaim and Murphy disclosures cited for 7.c and 7.g also show that claim 12 is obvious.

- xv. **[13] The system of claim 7, in which the healthcare data represents medical related information associated with the individuals of the group.**

110. Evenhaim describes or suggests that the healthcare data represents medical information associated with the individuals of the group. Ex. 1005, Abstract, Fig. 5 (showing that micro-data results are de-identified before being delivered), 10:46-54, 11:14-29, 12:36-47. Evenhaim also describes, in Figs 1 & 2, a prior art system in which the healthcare data is medical information associated with the individuals of a group. Figure 2 shows that medical data includes medical, hospital, and pharmacy claims data. *Id.*, 16:46-57, 17:27-41.

- xvi. **[1.a] A computer-implemented method comprising: receiving a request from a user device for information that characterizes a group composed of multiple individuals from a set of de-**

identified medical data representing characteristics of the individuals pre- or post-encounter with a healthcare facility,

111. In my opinion, Evenhaim in view of Murphy describes or suggests 1.a for the reasons cited in 7.a, 7.b, 9 above.

xvii. [1.b] wherein the request message for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,

112. In my opinion, Evenhaim in view of Murphy describes or suggests 1.b for the reasons cited in 7.a, 7.b above.

xviii. [1.c] and wherein the request message includes a batch number and at least one report option;

113. In my opinion, Evenhaim in view of Murphy describes or suggests 1.c for the reasons cited in 7.h above.

114. Further, it would be obvious to a POSA to use a batch number or other means to identify groups or programs in the requests and data and to aid the user in understanding the results. Evenhaim suggests including a batch number in the request message. For example, Evenhaim shows in Fig. 4A a request to analyze Health Programs 1 through n is made and that a study control group, e.g., HP(1-n) is defined. *Id.*, Fig. 4A. Evenhaim further explains that a group name is provided in the result of a query. *Id.*, 16:22-32. It would be obvious to a POSA that numbering health programs, defining study groups, or using group names could also be forms of or interchangeable with using a batch number or any other means of identifying

the groups or programs. A POSA would also include the group name or batch number in the request so the user could understand the results, which is suggested by Evenhaim. *See id.*, 10:37-54.

115. Further, a POSA would understand that to manage multiple queries and multiple data requests, one would need to implement a scheme to identify which query a request is related to, such as including a batch number. A POSA would understand that a server that processes multiple requests would assign batch numbers to those requests in order to be able to process them in the order they are received. This is a standard queuing technique that has been known and in use in computer programming since the first programmable computers were invented.

116. Evenhaim also suggests that the request can include a report option by permitting specific formats for answers to a query. *Id.*, 8:65-9:12 (“[T]he query asks to characterize patients who have a specific health problem and who receive a specific therapy in terms of the seniority of their attending physicians. . . . [T]he answer is a table of seniority brackets of physicians compared to a data cluster of multiple patients’ data.”).

117. A POSA would understand that in order to accommodate the different needs of the different requesters, one would need to allow the requester to specify how the information should be reported in the form of including a report option in the request.

- xix. [1.d] producing, by a processor of an encryption server, a unique request token from identifying data for each individual included in the group,**

118. In my opinion, Evenhaim in view of Murphy describes or suggests 1.d for the reasons cited in 7.b, 7.c, 7.e, and 11 above.

- xx. [1.e] wherein the request tokens are encrypted before being transmitted over a network;**

119. In my opinion, Evenhaim in view of Murphy describes or suggests 1.e for the reasons cited in 11, 7.a above.

- xxi. [1.f] comparing the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens;**

120. In my opinion, Evenhaim in view of Murphy describes or suggests 1.f for the reasons cited in 7.e above.

- xxii. [1.g] producing the healthcare information with the identity of the individuals being unattainable from the produced healthcare information**

121. In my opinion, Evenhaim in view of Murphy describes or suggests 1.g for the reasons cited in 7.f above.

- xxiii. [1.h] based on a minimum number of matched tokens; and**

122. In my opinion, Evenhaim in view of Murphy describes or suggests 1.h for the reasons cited in 7.k, and 7.l above.

- xxiv. [1.i] providing the produced healthcare information to report the characteristics of the group absent authorization from the individuals in the group.**

123. In my opinion, Evenhaim in view of Murphy describes or suggests 1.i for the reasons cited in 7.h above.

xxv. [2] The computer-implemented method of claim 1, wherein the provided healthcare information is in conformity with the Health Insurance Portability and Accountability Act (HIPAA).

124. In my opinion, Evenhaim in view of Murphy describes or suggests 2 for the reasons cited in 8 above.

xxvi. [3] The computer-implemented method of claim 1, wherein receiving the request for healthcare information that characterizes the group of individuals includes receiving personally identifiable information associated with the individuals of the group.

125. In my opinion, Evenhaim in view of Murphy describes or suggests 3 for the reasons cited in 9 above.

xxvii. [4] The computer-implemented method of claim 1, in which producing the requested healthcare information includes encrypting respective data that identifies each individual.

126. In my opinion, Evenhaim in view of Murphy describes or suggests 4 for the reasons cited in 10 above.

xxviii. [5] The computer-implemented method of claim 1, in which the tokens associated with healthcare data and the request tokens are similarly encrypted.

127. In my opinion, Evenhaim in view of Murphy describes or suggests 5 for the reasons cited in 11 above.

xxix. [6] The computer-implemented method of claim 1, in which the healthcare data represents medical related information associated with the individuals of the group.

128. In my opinion, Evenhaim in view of Murphy describes or suggests 6 for the reasons cited in 13 above.

xxx. [14.a] One or more computer readable storage devices storing instructions that are executable by a processing device, and upon such execution cause the processing device to perform operations comprising:

129. In my opinion, Evenhaim in view of Murphy describes or suggests 14.a for the reasons cited in 7.a above.

xxxi. [14.b] receiving a request from a user device for information that characterizes a group composed of multiple individuals from a set of de-identified data representing characteristics of the individuals pre-or post-encounter with a healthcare facility,

130. In my opinion, Evenhaim in view of Murphy describes or suggests 14.b for the reasons cited in 7.b above.

xxxii. [14.c] wherein the request message for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,

131. In my opinion, Evenhaim in view of Murphy describes or suggests 14.c for the reasons cited in 7.c above.

xxxiii. [14.d] and wherein the request message includes a batch number and at least one report option;

132. In my opinion, Evenhaim in view of Murphy describes or suggests
14.d for the reasons cited in 1.c above.

xxxiv. [14.e] producing, by a processor of an encryption server, a unique request token from identifying data for each individual included in the group,

133. In my opinion, Evenhaim in view of Murphy describes or suggests
14.e for the reasons cited in 7.d and 11 above.

xxxv. [14.f] wherein the request tokens are encrypted before being transmitted over a network;

134. In my opinion, Evenhaim in view of Murphy describes or suggests
14.f for the reasons cited in See 7.a and 11 above.

xxxvi. [14.g] comparing the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens;

135. In my opinion, Evenhaim in view of Murphy describes or suggests
14.g for the reasons cited in 7.e above.

xxxvii. [14.h] producing the healthcare information with the identity of the individuals being unattainable from the produced healthcare information

136. In my opinion, Evenhaim in view of Murphy describes or suggests
14.h for the reasons cited in 7.f and 7.g above.

xxxviii. [14.i] based on a minimum number of matched tokens; and

137. In my opinion, Evenhaim in view of Murphy describes or suggests
14.i for the reasons cited in 7.g above.

xxxix. [14.j] providing the produced healthcare information to report the characteristics of the group absent authorization from the individuals in the group.

138. In my opinion, Evenhaim in view of Murphy describes or suggests 15 for the reasons cited in 7.h and 7.i above.

i. [15] The computer readable storage devices of claim 14, wherein the provided information is medical data in conformity with the Health Insurance Portability and Accountability Act (HIPAA).

139. In my opinion, Evenhaim in view of Murphy describes or suggests 15 for the reasons cited in 8 above.

ii. [16] The computer readable storage devices of claim 14, wherein receiving the request for healthcare information that characterizes the group of individuals includes receiving personally identifiable information associated with the individuals of the group.

140. In my opinion, Evenhaim in view of Murphy describes or suggests 16 for the reasons cited in 9 above.

i. [17] The computer readable storage devices of claim 14, in which producing the requested healthcare information includes encrypting respective data that identifies each individual.

141. In my opinion, Evenhaim in view of Murphy describes or suggests 17 for the reasons cited in 10 above.

ii. [18] The computer readable storage devices of claim 14, in which the tokens associated with the healthcare data and the request tokens are similarly encrypted.

142. In my opinion, Evenhaim in view of Murphy describes or suggests 18 for the reasons cited in 11 above.

- iii. **[19] The computer readable storage devices of claim 14, in which the healthcare data represents medical related information associated with the individuals of the group.**

143. In my opinion, Evenhaim in view of Murphy describes or suggests 19 for the reasons cited in 13 above.

- iv. **[20] The computer readable storage devices of claim 14, in which producing the requested healthcare information includes determining if the group includes at least the minimum number of individuals.**

144. In my opinion, Evenhaim in view of Murphy describes or suggests 19 for the reasons cited in 7.c above.

- v. **[21] The computer-implemented method of claim 1, wherein the identity of the individuals is unattainable from the unique request tokens.**

145. In my opinion, Evenhaim in view of Murphy describes or suggests 20 for the reasons cited in 7.f and 7.g above.

- vi. **[22] The computer-implemented method of claim 1, wherein the unique request tokens and the tokens associated with the healthcare data are similarly encrypted.**

146. In my opinion, Evenhaim in view of Murphy describes or suggests 20 for the reasons cited in 11 above.

C. Ground 2: Obvious (§ 103) over Evenhaim in view of Settimi and Murphy.

147. In my opinion, each element of the '375 patent's claims 1-22 is obvious to a POSA over Evenhaim in view of Settimi and Murphy. Support for my opinions is detailed further below.

148. A POSA would look to combine these three references because they seek to solve the same problem of providing access to patient data in a manner that is compliant with privacy laws such, as HIPAA. In my opinion, a POSA would look to combine Evenhaim's distribution system and ability to query healthcare data with the centralized server of de-identified patient data taught by Settimi to streamline the process of querying the data sources in Evenhaim. A POSA would have been further motivated to combine Evenhaim and Settimi with Murphy to ensure that the returned results could not be sufficiently small for a user to be able to identify an individual.

149. A POSA would find that Settimi's centralized warehouse of de-identified patient data would be a more convenient and reliable way to update data and obtain complete responses to the query for healthcare information. Settimi describes creating encrypted tokens, called EPIDs, and inserting them into patient records to replace personally identify information. Ex. 1006, [0039]-[0041] ("methods of creating the EPID . . . may be [based on] patient name, age and social security number, etc."), [0042]. Settimi describes that whenever patient records are

updated at the source, for example in a doctor's office, they are de-identified with the encrypted tokens (EPIDs) and transmitted to a centralized data warehouse where they can be queried in compliance with HIPAA. Id., [0037]. The EPIDs can then be used to query the de-identified patient data from a single source and match records that belong to the same patient while also protecting the patient's identity in compliance with HIPAA. Id., [0037], [0043].

150. While Evenhaim sends the queries for healthcare data to various healthcare organizations, health plans (e.g., HMOs), and pharmacy benefit managers, etc. Settimi provides the added advantage of consolidation information from healthcare sources into a central repository for queries. See Ex. 1005, 15:15-24, 16:5-21, Fig. 4; Ex. 1006, Figs. 1 & 2, [0027]. The repository in Settimi provides one central database for conducting searches and also uses a "common vocabulary" that makes searching easier and more reliable (e.g., the searcher does not need to know the proper search terms used by different plans/providers). Ex. 1006, [0009], [0060]-[0062]. Further, it overcomes the prior central warehouse designs in that, as discussed in more detail below, common encryption allows for search for specific patients in the central database, even though there has already been de-identification. A POSA would further look to Murphy's algorithm because Murphy provides additional information about protecting the identity of individuals by preventing their identity from being revealed when a parameter-

based search is performed on a medical database in compliance with HIPAA. This aligns with one of the goals of Dick and Landi. More specifically, Murphy more fully explores the security issues related trying to identifying a patient based the results of query that receives de-identified medical information, something a POSA implementing the invention in Dick in view of Landi would be motivated to use in coming up with a secure system that complied with HIPAA and other privacy regulations.

151. Murphy describes the results of a study which sought to prevent the identification of individuals in a search on a medical database. Ex. 1009, p. 553. One way Murphy did this was by ensuring that more than one individual is returned when a search is performed. Id., p. 555. Murphy describes an algorithm which blocks queries that return fewer than 3 patients thereby preventing disclosing the patient's identity. Id., p. 555.

152. Evenhaim and Settimi can be combined with a reasonable expectation of success because they both are performed functions in the same technical field. For example, both operate in a database distribution network to provide access to healthcare information in a manner that complies with HIPAA and obviates the need to obtain explicit patient permission. Also, they both use similar processes of de-identifying data by creating encrypted tokens to replace personal identifying information in a manner that leaves it in usable form. Both describe encrypted

tokens that are created in a similar fashion from person identifying information. A POSA would have a reasonable expectation of success in combining Murphy with Evenhaim and Settimi because preventing the return of a query of fewer than three patient records would have been a simple programming task in systems already designed to return search results. Moreover, as discussed above, A POSA would find benefit in applying Settimi's central database to Evenhaim to streamline, simplify and improve the querying process. Because of these similarities and the obvious benefits of applying Settimi's central database, a POSA would be motivated to combine these references with a reasonable expectation of success. A POSA would have a reasonable expectation of further combining with Murphy because all three are directed to searching for sensitive information in a database and providing that information in a way that an individual cannot be identified. Evenhaim, Settimi and Murphy are performed in similar computing environments so a POSA would be able to combine the references with a reasonable expectation of success. Moreover, making a modification according to what Murphy describes, i.e., ensuring a minimum number of results is returned, is a simple design choice for a POSA to make when seeking access to the healthcare information.

i. [7.a] A system comprising: a data server device

153. Per my discussion for element 7.a in Ground 1, Evenhaim in view of Murphy describes this element at the relevant time period.

ii. [7.b] for producing healthcare information that characterizes a group of a number of individuals identified in a request for the healthcare information that characterizes the group

154. Per my discussion for element 7.b in Ground 1, Evenhaim in view of Murphy describes this element at the relevant time period.

155. In the combination with Settimi, the produced tokens created in step 3 in Fig. 4 of Evenhaim would be used to query a central warehouse in Settimi rather than being sent to each individual source. Therefore, Health Plan/PBM 1 through n shown in Fig. 4 of Evenhaim would de-identify its patient records whenever those records were updated and send the updated records to the central warehouse described in Settimi. Under this combination the data warehouse would replace the Health Plan/PBM's data source to be queried.

156. Because the central warehouse only contains de-identified records which have encrypted tokens in them in place of personal identifying information, a POSA would understand that it is necessary that the encrypted patient list sent in Evenhaim in steps 3-4 would be converted to encrypted tokens (the produced tokens) before being sent the central warehouse (the data source) for querying. Ex. 1005, Fig. 4, 12:4-13:20. Using the same or similar public encryption key to create the produced tokens as was used to create the corresponding tokens in the de-identified patient records allows the records to be matched. *Id.*

157. In Evenhaim the list of individuals in the group of people to be used for the query (shown in box 2 of Fig. 4) is used to create produced tokens to be used for the query as described in the specification. *Id.*, Fig. 4, 12:4-13:20. “[T]he instant method can ask about the status of Tom, Dick, Harry, and a further collection of individuals – and obtain results from that query – so long as the statistical character of those results will not allow a correlation of a result back to an individual.”). Both Evenhaim and Settimi propose using encrypted tokens in place of personal identifying information to accomplish the privacy goal of the inventions. In Settimi, EPIDs are produced for each individual. A POSA would understand that the EPIDs are tokens. Ex. 1006, [0039]-[0042].

158. In box 4 of Fig. 4, the query would contain the produced tokens that were made from the encrypting the list of individuals in the group that was provided for the query. Ex. 1005, Fig. 4, 7. It also would have been obvious to make a request for healthcare information in view of Settimi. For example, Settimi describes being able to query the de-identified healthcare information. Ex. 1006, [0035] (“The de-identified patient data is then staged into a data warehouse 210 where it is available for querying.”), [0036]. It would have been obvious to a POSA to request information for a group of individuals with certain characteristics in view of Settimi’s description of collecting this information.

- iii. **[7.c] wherein the request for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,**

159. Per my discussion for element 7.c in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- iv. **[7.d] wherein a unique request token is produced for each individual included in the group**

160. Per my discussion for element 7.d in Ground 1, Evenhaim in view of Murphy describes and/suggests this element at the relevant time period.

161. In the combination with Settimi, the produced tokens created in step 3 in Fig. 4 of Evenhaim would be used to query a central warehouse in Settimi rather than being sent to each individual source. Therefore, Health Plan/PBM 1 through n shown in Fig. 4 of Evenhaim would de-identify its patient records whenever those records were updated and send the updated records to the central warehouse described in Settimi. Under this combination the data warehouse would replace the Health Plan/PBM's data source to be queried.

162. Because the central warehouse only contains de-identified records which have encrypted tokens in them in place of personal identifying information, a POSA would understand that it is necessary that the encrypted patient list sent in Evenhaim in steps 3-4 would be converted to encrypted tokens (the produced tokens) before being sent the central warehouse (the data source) for querying. Ex.

1005, Fig. 4, 12:4-13:20. Using the same or similar public encryption key to create the produced tokens as was used to create the corresponding tokens in the de-identified patient records allows the records to be matched. *Id.*

163. In Evenhaim the list of individuals in the group of people to be used for the query (shown in box 2 of Fig. 4) is used to create produced tokens to be used for the query as described in the specification. *Id.*, Fig. 4, 12:4-13:20. “[T]he instant method can ask about the status of Tom, Dick, Harry, and a further collection of individuals – and obtain results from that query – so long as the statistical character of those results will not allow a correlation of a result back to an individual.”). Both Evenhaim and Settimi propose using encrypted tokens in place of personal identifying information to accomplish the privacy goal of the inventions. In Settimi, EPIDs are produced for each individual. A POSA would understand that the EPIDs are tokens. Ex. 1006, [0039]-[0042].

164. In box 4 of Fig. 4, the query would contain the produced tokens that were made from the encrypting the list of individuals in the group that was provided for the query. Ex. 1005, Fig. 4, 7. It also would have been obvious to make a request for healthcare information in view of Settimi. For example, Settimi describes being able to query the de-identified healthcare information. Ex. 1006, [0035] (“The de-identified patient data is then staged into a data warehouse 210 where it is available for querying.”), [0036]. It would have been obvious to a

POSA to request information for a group of individuals with certain characteristics in view of Settimi's description of collecting this information.

- v. **[7.e] the data server is configured to compare the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens,**

165. Per my discussion for element 7.e in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

166. Settimi also describes or suggests that each token in the query is associated with a corresponding token in the de-identified healthcare information in the data warehouse. Ex. 1006, [0027], [0030], [0035] ("The de-identified patient data is then staged into a data warehouse 210 where it is available for querying."), [0036], [0037] ("create[ing] patient cohort reports 212 based on the querying the longitudinal patient data is supported by the ability to connect all records relating to a single patient in the data warehouse 210. This requires a unique identifier to be associated with each patient record that is transmitted to the data warehouse 210. The unique identifier indicates an anonymous or abstract patient having certain characteristics but does not provide personal direct identifying information such as name, social security number, street address, etc."), [0039-42].

167. As shown in Figs. 4 and 4A, the produced tokens in the query are sent to the data server (boxes 3-5), which in this combination would be the centralized warehouse de-identified data of Settimi. In both Evenhaim and Settimi, de-

identification is done by substituting encrypted tokens into the patient records to replace identifying information. Ex. 1006, Figs. 1 and 2, [0027], [0037]-[0043].

168. When the data server receives the produced tokens in the query, those tokens are matched against corresponding tokens in patient data. Ex. 1005, 12:4-47. In this combination the data server is the centralized warehouse containing only de-identified patient data with encrypted tokens, thus a POSA would readily understand that the search of patient records involves matching the produced tokens from the query against the corresponding tokens in the patient records. No name or other identifying information is present in the patient records and therefore having produced tokens is necessary to query the patient data for the group of individuals in the Evenhaim query. Additionally, Settini also describes this type of search. Ex. 1006, [0037] (“This requires a unique identifier to be associated with each patient record that is transmitted to the data warehouse 210. The unique identifier indicates an anonymous or abstract patient”). Thus, the tokens sent from the aggregator are matched against the tokens in the de-identified patient records in the data source to retrieve the de-identified patient data. Ex. 1005, 12:13-16, 12:42-45.

169. A POSA would understand that the token creation and matching in Evenhaim, when combined with the central data warehouse of Settini, would apply to the embodiment in Fig. 4 and 4A. Therefore, the data source in box 6 (the

central data warehouse) would have de-identified records with encrypted tokens that are identified in response to a query to the data sources sent in box 4. As taught by Evenhaim, as shown in Fig. 4, the tokens would be created on the aggregator using the same or similar public encryption key in order to allow the patient records to be matched. *Id.*, Fig. 4, 12:4-13:20. The specification describes that the token creation could be performed on the aggregator as well as on the data sources. *Id.*, 12:26-34.

vi. [7.f] the produced healthcare information is absent personally identifiable information and

170. Per my discussion for element 7.f in Ground 1, Evenhaim in view of Murphy sufficiently describes or suggests this element at the relevant time period.

vii. [7.g] is based on a minimum number of matched tokens,

171. Per my discussion for element 7.g in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

viii. [7.h] the data server is also configured to provide the produced healthcare information to report the characteristics of the group

172. Per my discussion for element 7.h in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

ix. [7.i] absent authorization from the individuals in the group.

173. Per my discussion for element 7.i in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- x. **[8] The system of claim 7, wherein the provided healthcare information is in conformity with the Health Insurance Portability and Accountability Act (HIPAA).**

174. Per my discussion for element 8 in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xi. **[9] he system of claim 7, wherein the request for healthcare information that characterizes a group of individuals includes personally identifiable information associated with the individuals of the group.**

175. Per my discussion for element 9 in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xii. **[10] The system of claim 7, in which the data that identifies each individual is encrypted.**

176. Per my discussion for element 10 in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xiii. **[11] The system of claim 7, in which the tokens associated with the healthcare data and the request tokens are similarly encrypted.**

177. Per my discussion for element 11 in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xiv. **[12] The system of claim 7, in which the data server provides the requested healthcare information if the group includes at least the minimum number of individuals.**

178. Per my discussion for element 12 in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xv. **[13] The system of claim 7, in which the healthcare data represents medical related information associated with the individuals of the group.**

179. Per my discussion for element 13 in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xvi. **[1.a] A computer-implemented method comprising: receiving a request message for healthcare information that characterizes a group composed of a number of individuals,**

180. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.a for the reasons cited in 7.a, 7.b, 9 above.

- xvii. **[1.b] wherein the request message for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,**

181. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.b for the reasons cited in 7.a, 7.b above.

- xviii. **[1.c] and wherein the request message includes a batch number and at least one report option;**

182. Per my discussion for element 1.c in Ground 1, Evenhaim in view of Murphy describes or suggests this element at the relevant time period.

- xix. **[1.d] producing, by a processor of an encryption server, a unique request token from identifying data for each individual included in the group,**

183. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.d for the reasons cited in 7.b, 7.c, 7.e, and 11 above.

- xx. [1.e] wherein the request tokens are encrypted before being transmitted over a network;**

184. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.e for the reasons cited in 11, 7.a above.

- xxi. [1.f] comparing the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens;**

185. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.f for the reasons cited in 7.e above.

- xxii. [1.g] producing the healthcare information with the identity of the individuals being unattainable from the produced healthcare information**

186. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.g for the reasons cited in 7.f above.

- xxiii. [1.h] based on a minimum number of matched tokens; and**

187. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.h for the reasons cited in 7.g above.

- i. [1.i] providing the produced healthcare information to report the characteristics of the group absent authorization from the individuals in the group.**

188. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 1.h for the reasons cited in 7.h above.

- ii. [2] The computer-implemented method of claim 1, wherein the provided healthcare information is in conformity with the Health Insurance Portability and Accountability Act (HIPAA).**

189. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 2 for the reasons cited in 9 above.

- iii. **[3] The computer-implemented method of claim 1, wherein receiving the request for healthcare information that characterizes the group of individuals includes receiving personally identifiable information associated with the individuals of the group.**

190. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 3 for the reasons cited in 10 above.

- iv. **[4] The computer-implemented method of claim 1, in which producing the requested healthcare information includes encrypting respective data that identifies each individual.**

191. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 4 for the reasons cited in 11 above.

- v. **[5] The computer-implemented method of claim 1, in which the tokens associated with healthcare data and the request tokens are similarly encrypted.**

192. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 5 for the reasons cited in 13 above.

- vi. **[6] The computer-implemented method of claim 1, in which the healthcare data represents medical related information associated with the individuals of the group.**

193. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 6 for the reasons cited in 7.a above.

- vii. **[14.a] One or more computer readable storage devices storing instructions that are executable by a processing device, and**

upon such execution cause the processing device to perform operations comprising:

194. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.a for the reasons cited in 7.a above.

- viii. [14.b] receiving a request from a user device for information that characterizes a group composed of multiple individuals from a set of de-identified data representing characteristics of the individuals pre-or post-encounter with a healthcare facility,**

195. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.b for the reasons cited in 7.b above.

- ix. [14.c] wherein the request message for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,**

196. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.c for the reasons cited in 7.c above.

- x. [14.d] and wherein the request message includes a batch number and at least one report option;**

197. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.d for the reasons cited in 1.c above.

- xi. [14.e] producing, by a processor of an encryption server, a unique request token from identifying data for each individual included in the group,**

198. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.e for the reasons cited in 7.d, 11 above.

- xii. **[14.f] to determine whether the group includes at least a minimum number of individuals to be included in a report in response to the request,**

199. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.f for the reasons cited in 7.a, 11 above.

- xiii. **[14.g] comparing the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens;**

200. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.g for the reasons cited in 7.e above.

- xiv. **[14.h] producing the healthcare information with the identity of the individuals being unattainable from the produced healthcare information**

201. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.h for the reasons cited in 7.f, 7.g above.

- xv. **[14.i] based on a minimum number of matched tokens; and**

202. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.i for the reasons cited in 7.g above.

- xvi. **[14.j] providing the produced healthcare information to report the characteristics of the group absent authorization from the individuals in the group.**

203. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 14.i for the reasons cited in 7.h, 7.i above.

- xvii. **[15] The computer readable storage devices of claim 14, wherein the provided healthcare information is in conformity**

with the Health Insurance Portability and Accountability Act (HIPAA).

204. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 15 for the reasons cited in 8 above.

xviii. [16] The computer readable storage devices of claim 14, wherein receiving the request for healthcare information that characterizes the group of individuals includes receiving personally identifiable information associated with the individuals of the group.

205. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 16 for the reasons cited in 9 above.

xix. [17] The computer readable storage devices of claim 14, in which producing the requested healthcare information includes encrypting respective data that identifies each individual.

206. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 17 for the reasons cited in 10 above.

xx. [18] The computer readable storage devices of claim 14, in which the tokens associated with the healthcare data and the request tokens are similarly encrypted.

207. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 18 for the reasons cited in 11 above.

xxi. [19] The computer readable storage devices of claim 14, in which the healthcare data represents medical related information associated with the individuals of the group.

208. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 19 for the reasons cited in 13 above.

- xxii. [20] The computer readable storage devices of claim 14, in which producing the requested healthcare information includes determining if the group includes at least the minimum number of individuals.**

209. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 20 for the reasons cited in 7.c above.

- xxiii. [21] The computer-implemented method of claim 1, wherein the identity of the individuals is unattainable from the unique request tokens.**

210. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 21 for the reasons cited in 7.f, 7.g above.

- xxiv. [22] The computer-implemented method of claim 1, wherein the unique request tokens and the tokens associated with the healthcare data are similarly encrypted.**

211. In my opinion, Evenhaim in view of Settimi and Murphy describes or suggests 22 for the reasons cited in 11 above.

D. Ground 3: Obvious (§ 103) over Dick in view of Landi and Murphy

212. In my opinion, each element of the '375 patent's claims 1-20 is obvious to a POSA over Dick in view of Landi and in further view of Murphy. Support for my opinions is detailed further below.

213. Dick describes a system for requesting and retrieving de-identified medical information for a group of individuals using a first server, called the Requester's Server, to send a query to a data server, referred to as a medical

repository, which searches for matching patient records and returns de-identified healthcare information in response to the query. Ex. 1007, Abstract; claim 23.

Dick's system produces healthcare information for many purposes, including in connection with clinical trials for new pharmaceuticals, as well as for risk assessment for health insurance underwriting. *Id.*, [0003], [0026], [0037], [0041]-[0042], [0047].

214. Dick describes or suggests de-identification for complying with HIPAA. It describes that de-identification can take place at the medical repository or originating source of the information, and that a POSA would understand it could use "de-identifying methods and processes," but Dick does not describe any specific method to use for this de-identification. *Id.*, [0013], [0046]. Therefore, a POSA relying on Dick would look to other sources to implement the de-identification called for in Dick.

215. Landi sets forth in detail a de-identification method that a POSA would understand can be implemented in Dick. Landi describes a token matching system used to de-identify patient records to comply with HIPAA. More specifically, Landi describes how healthcare providers or sources can de-identify their records by creating an encrypted token, then release the de-identified records to a central data server where they can be processed, e.g., queried for research and healthcare monitoring. Ex. 1008, 4:19-29.

216. A POSA would be motivated to combine the tokenization and de-identified central data server teachings of Landi to Dick's system for querying de-identified healthcare information that characterizes a group of individuals in order to providing a HIPAA compliant search for healthcare information. Dick describes using de-identification in its system but does not specify any particular method for doing that, and therefore a POSA would look to Landi's use of an Encrypted ID (encrypted token) to de-identify the results of a search because both methods relate to queries of medical records to produce de-identified results in compliance with HIPAA. So, for example, a POSA would understand that insurance risk scoring queries described in Dick could implement the token creation and linking process described in Landi and use the centralized data server in Landi to simplify the querying of the de-identified patient records.

217. It would be obvious to a POSA that to use the public key from Landi to create a list of Encrypted IDs to query the central Landi's central database of de-identified medical records using Dick's system for searching a central database for a group of individuals. *See* Ex. 1007, [0011], [0024]-[0025]. This would be compliant with the purpose of the Encrypted IDs in Landi, and allow each patient to be recognized without revealing the identity of the patients. Ex. 1008, 8:38-55.

218. A POSA would have also understood that when using Dick's system for searching for a group of individuals in Landi's central database, querying the database with something other than the Encrypted IDs would not return proper results. Therefore, it would be obvious to a POSA to use Landi's public key in Dick's system for searching. This would be supported by Dick's system for searching that would use any "identification characteristics capable of being correlated with stored medical information." Ex. 1007, [0024].

219. A POSA would have a reasonable expectation of success combining the two references because both provide systems that search and evaluate a collection of medical information and seek to provide it in a manner that does not reveal the identity of individuals and is still useful to the requester. Landi's detailed disclosure on how to create encrypted tokens from identifying information using the same or similar public encryption keys fits perfectly into Dick's system for producing de-identified healthcare information that characterizes a group of individuals, and would work well for accomplishing the examples described in Dick, such as querying for healthcare information based on identifying information in order to do risk analysis for health insurance, while complying with HIPAA. Additionally, a POSA would know that the token creation and matching system described in Landi was standard in the industry at

the time, and therefore would have reasonably expected to be able to successfully implement such a scheme.

220. A POSA would then look to combine Dick and Landi with Murphy because Murphy provides additional information about protecting the identity of individuals by preventing their identity from being revealed when a parameter-based search is performed on a medical database in compliance with HIPAA. This aligns with one of the goals of Dick and Landi. More specifically, Murphy more fully explores the security issues related trying to identifying a patient based the results of query that receives de-identified medical information, something a POSA implementing the invention in Dick in view of Landi would be motivated to use in coming up with a secure system that complied with HIPAA and other privacy regulations.

221. Murphy describes the results of a study which sought to prevent the identification of individuals in a search on a medical database. Ex. 1009, p. 553. One way Murphy did this was by ensuring that more than one individual is returned when a search is performed. Id., p. 555. Murphy describes an algorithm which blocks queries that return fewer than 3 patients thereby preventing disclosing the patient's identity. Id., p. 555.

222. A POSA would have a reasonable expectation of success combining Dick, Landi and Murphy because they are all directed to searching for sensitive

information in a database and providing that information in a way that an individual cannot be identified. Dick, Landi and Murphy are performed in similar computing environments so a POSA would be able to combine the references with a reasonable expectation of success. Moreover, making a modification according to what Murphy describes, i.e., ensuring a minimum number of results is returned, is a simple design choice for a POSA to make when seeking access to the healthcare information.

xxv. [7.a] A system comprising: a data server device

223. Dick describes a data server, called a Medical Information Repository. Ex. 1007, Fig. 3, [0019], [0028] (“Following receipt of the request for information 105, the Central Server 245 will transmit a query to a medical information repository 275 for information pursuant to the request 115.”). In addition, as shown in FIG. 3, an Archive Medical Information System (AMIS) server 265 may be accessed by the central server 245. *Id.*, Fig. 3, [0028]-[0031]. Dick even acknowledges that it would be apparent to a POSA that that its system “may be implemented by programming one or more suitable general-purpose computers having appropriate hardware.” *Id.*, [0048].

xxvi. [7.b] for producing healthcare information that characterizes a group of a number of individuals identified in a request for the healthcare information that characterizes the group

224. Dick explains that its data server, or Medical Information Repository, produces the healthcare information that characterizes a group of individuals. Ex. 1007, Fig. 3, [0019], [0028] (“Following receipt of the request for information 105, the Central Server 245 will transmit a query to a medical information repository 275 for information pursuant to the request 115.”). As shown in FIG. 3, an Archive Medical Information System (AMIS) server 265 may be accessed by the central server 245. *Id.*, Fig. 3, [0028]-[0031]. AMIS server 265 will receive the transmitted query 115 and proceed to search the AMIS database for information in satisfaction of the query. *Id.* Dick further explains that the request for medical records is sent to a repository to obtain healthcare information for purposes including for risk assessment of the individuals for policy underwriting. Ex. 1007, [0003], [0011], [0019], [0026], [0037], [0041]-[0042], [0047].

225. As shown in claim 23, Dick’s Medical Information Repository distributes the healthcare information in response to queries. Ex. 1007, claim 23 (“(b) compiling and formatting at the site of the medical information repository the results of the search in a generic format ascertainable by requester;...(d) providing the compiled and de-identified results of at least one of the preselected individuals to a requester”).

226. A POSA would implement Landi’s token system and central database of de-identified healthcare data into Dick’s system for producing healthcare

information. Per Landi, the system operates at different clinical sites and utilizes central patient data processing system to collect de-identified data records. *Id.*, 4:19-28; 5:40-44, 14:26-32. A POSA would have understood that a data server could produce the requested healthcare data because it was a common element in computer systems operating across multiple sites. Additionally, a POSA would have known that the central patient data processing system in Landi, which collects de-identified patient records, would work to provide de-identified healthcare data in response to the query from Dick's Requester's Server.

xxvii. [7.c] wherein the request for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,

227. Dick describes a "Requester's Server" that receives the names of a group of individuals from a "Requesting & Viewing Client," and then creates a query that it transmits to a "Medical Information Repository." Ex. 1007, Fig. 3, [0048], claim 23. This system can be used to query for de-identified healthcare information to comply with HIPAA but does not mandate any specific method for de-identifying the information. *Id.*, [0044]-[0046]. Therefore, a POSA would look to Landi as a known option for the de-identification called for in Dick. In doing so a POSA would be motivated to incorporate Landi's encryption server into Dick's Requester Server to create encrypted tokens, which would be used to query a de-identified data server as discussed further below.

228. Landi describes a system comprising an encryption server (called an encryption system) that is used to create tokens from personally identifying information using a public encryption key. Ex. 1008, 2:63-67, 3:3-5, 14:29-32. A POSA implementing Dick's system for querying medical information would use Landi's encryption server to query the data servers for medical information. *Id.*, 8:47-55. Landi makes clear that its encryption server may be implemented in any number of hardware, software, and firmware combinations. *Id.*, 3:58-4:2.

229. Further, a server capable of performing encryption was well-known elements at the time of invention.

230. The encryption server being configured to determine if the number of individuals in a request is at least a minimum number of individuals would have been obvious to a POSA as a further level of protection to comply with HIPAA regulations. It would have been a simple design choice to ensure that a minimum number of individuals were identified in the request. If only one individual was identified, then there is a significant risk that an individual could be identified based on the selected parameters.

231. Landi describes or suggests identifying at least a minimum number of individuals in a request because Landi is intended to comply with the provisions of HIPAA. Ex. 1008, 2:5-18; 45 C.F.R. § 164.514(b) ("a person with appropriate knowledge and experience with generally accepted statistical and scientific

principles and methods for rendering information not individually identifiable”). Similarly, Dick cites to complying with HIPAA using the statistical de-identification method, which a POSA would understand to include optionally using a minimum number of individuals in a request. Ex. 1007, [0045] (“Where removing such information is not practical or renders the data useless, the medical information repository may instead release medical information upon condition that a person with appropriate knowledge and experience with generally accepted statistical and scientific principles has applied a ‘de-identifying’ methodology to remove or make certain that no personally identifiable information is released from the repository.”), [0046].

232. A POSA would understand that one way to ensure compliance with 45 C.F.R. § 164.514(b) would be to restrict the query for patient information to at least a minimum number of patients so that information returned could not be associated with a specific individual. 45 C.F.R. § 164.514(b), which governs the use and disclosure of protected information, provides that “a person with appropriate knowledge and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable” can apply those principles to determine that the “risk is very small that the information could be used, alone or in combination with other reasonably available information” to identify the patient. If a search only requests one

individual, or a very small number of individuals, then it would not be HIPAA compliant because the risk would not be very small that the information could be used to identify the patient.

233. As discussed above, data servers such as those in Dick and Landi were known to be vulnerable to attacks designed to return greater a minimum number of individuals, but use targeted specific requests, in combination with known information, in order to determine the individuals' identities. Therefore, it would have been obvious to determine whether at least a minimum number of individuals have been included in a request in order to protect against such an attack. This would have been a simple design choice in order to comply with HIPAA.

xxviii. [7.d] wherein a unique request token is produced for each individual included in the group

234. Dick's Requester's Server receives a group of individuals and creates a query for those individuals to access corresponding de-identified healthcare information for that group. Ex. 1007, Figs. 1-3, [0044]-[0046], claim 23. The request for information in Dick includes the identification of a subject, which can be "a human individual or a group of humans." *Id.*, [0024]; *see also* Fig. 1, claim 23. Dick does not set forth a specific de-identification method so a POSA would understand to use a known method, such as Landi's token creation and matching system which includes a centralized server of de-identified patient records. Using

this de-identification scheme, Dick's Requester's Server would produce encrypted tokens from the personal identifying information in the group of preselected individuals.

235. Landi describes a token matching system that creates tokens called Encrypted IDs, for each individual. Ex. 1008, 4:40-46, 11:23-29. By using the same or similar public encryption key, the produced tokens in Landi will match de-identified patient records with corresponding tokens for the same patient, without revealing the identity of the patient. *Id.*, Fig. 1, elements 22-n and 52, 3:52-57, 4:40-46, 5:19, 5:32-44, 5:57-61, and claim 23.

236. Landi further describes that each original source of patient data, such as a doctor's office or hospital, referred to as patient data processing systems, releases de-identified patient data to a central data processing system where it can be queried. *Id.*, 4:11-28, 7:3-10, 9:42-60. Landi, like Dick, also describes or suggests producing a token for each of these individuals in response to a request to collect their healthcare information. *Id.*

237. Dick also describes that the query from the Requester's Server is for healthcare information that characterize the group of individuals. Ex. 1007, Abstract, [0003], [0024], [0026], [0037], [0041]-[0042], [0047]. Landi similarly describes how the central repository can be queried in this way. Ex. 1008, 7:3-10, 8:64-9:19, 9:43-10:14. A POSA applying the token scheme in Landi to Dick's

system for querying healthcare information would use the encryption server in Landi to create tokens for each individual in the group using the same key that was used to de-identify the patient records stored in the central repository of patient data thereby allowing the tokens in the repository to be matched to corresponding tokens in the de-identified records. *Id.*, 8:49-52 (“[T]he public keys can be used to produce a unique, reproducible encrypted version of patient identifying information. Since the encryption is reproducible, the same patient can be recognized at different times . . .”). Because the central repository in Landi only has de-identified patient records in it, a POSA would implement Landi’s encryption server into Dick’s Requester’s Server to create the produced tokens from the personal identifying information specified in the query. *Id.*, 7:3-10. Further, a POSA would find understand that Landi’s central database/system, 30 could replace Dick’s central server 245. Both the central database and central server act as central intermediaries between the client/requester and the data servers.

238. Additionally, a POSA would be motivated to implement Landi’s encryption server into Dick’s Requester’s Server because it would be preferred that the Request sent to the data server should not have personally identifying information. Dick seeks to provide access to de-identified healthcare information in a manner that complies with HIPAA, as an alternative to getting explicit patient

authorization. Without authorization, it could violate HIPAA to just send PII to be used in querying data records. So, in order to avoid potentially violating HIPAA, a POSA would have used a public encryption key to create tokens on a requesting server. Landi further provides details on how the token creation and matching system works, which a POSA would understand would be an obvious way to implement the de-identification called for by Dick.

239. A POSA would therefore be motivated to combine Dick and Landi since Landi's token matching system for de-identification complements the system described in Dick for querying de-identified healthcare information for a group of individuals. A POSA would also have a reasonable expectation of success in combining Dick and Landi, because both references in the same field (sharing de-identified healthcare information) using the well-known technique of creating encrypted tokens to replace identity information in patient records in order to comply with HIPAA.

xxix. [7.e] the data server is configured to compare the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens,

240. The identity of the individuals that correspond to the tokens (Encrypted IDs) described in Landi, and implemented on Dick's Requester Sever, are not obtainable from the tokens absent a private key which must be used to decrypt the identifying data and is described as only being accessible to an

authorized party. Ex. 1008, 7:22-30, 11:13-17. The private key would normally not be available during the querying process in order to comply with HIPAA, and only the originating source of the information, such as a hospital, or the government might have access to the private key, so for the purposes of this system combining Dick with Landi, the identity of the individuals is unattainable from the tokens. *Id.*, 7:22-30 (“it is assumed that the entity (or entities) that operate the central patient data processing system (30) does not have the authority to re-identify subject patients that are associated with the de-identified data records (35). In other words, the central patient data processing system (30) does not include an encryption system that is capable of decrypting Encrypted IDs that are stored in or with the de-identified patient data records (35).”), *Id.*, 6:61-65; 8:51-55 (“Since the encryption is reproducible [by applying the public encryption key], the same patient can be recognized at different times . . . **even though the identity of the person remains unknown.**”) (emphasis added).

241. Dick in view of Landi also describes or suggests that each produced token can be associated with a corresponding token in a set of de-identified healthcare data stored in the central data repository, the Medical Information Repository. Since central repository described in Landi only has de-identified patient records, a POSA would understand that Dick’s Requester’s Server will produce encrypted tokens from the identifying information provided in the query,

by using a public encryption key. *Id.*, 7:3-10. It is not possible to directly use the names and other identifying information in the query in Dick to match against the de-identified patient records in the central repository because those records do not have personal identifying information. They only have the encrypted tokens, so to match against them a POSA would apply Landi's teachings to Dick. A POSA would therefore implement Landi's encryption server into Dick's Requester's Server to produce tokens from the identifying information in the query to be sent to the central repository of de-identified patient records that have corresponding tokens in them that were created using the same or similar public encryption key. In this way, the produced tokens in query can be used to match to the corresponding tokens in patient records. *Id.*, 8:47-55 ("The public keys are freely distributable (no HIPAA issues) to healthcare organizations for de-identification of patient data. Moreover, the public keys can be used to produce a unique, reproducible encryption version of patient identifying information. Since the encryption is reproducible, the same patient can be recognized at different times . . . even though the identity of the person remains unknown.").

242. Landi describes or describes that the healthcare records in the central database, the Medical Information Repository in Dick, has tokens (Encrypted IDs) in place of the identifying information. Therefore, each corresponding token (that corresponds to the produced tokens in the query) is associated with de-identified

healthcare records for an individual patient. HIPAA is concerned with protecting individuals' privacy as it relates to their healthcare data. *Id.*, 1:58-65, 3:1-5 (“[T]he encrypted ID comprises an encrypted representation of one or more items of individual identifying information, and wherein the encryption system can generate de-identified data records of the individual[s] which are associated with the encrypted ID.”).

243. As Landi makes clear throughout, it is concerned with protecting patient medical data. A POSA would therefore be motivated to use Landi's encrypted token scheme in Dick's system for producing healthcare information for a group of individuals in order to comply with HIPAA.

244. Landi also describes how this token matching system would work for querying for patient records associated with a clinical trial where “each time data is collected, the patient should be de-identified in the same way” in order to create a first token to be used to query the patient records. Ex. 1008, 9:43-60. Landi also other examples of querying the de-identified patient data. *Id.* 7:3-10; 8:47-55; 10:52-63; 11:23-29; claim 15 (“using the encrypted ID . . . to recognize a subject patient of patient data records collected at different times.”). Therefore, a POSA implementing Landi's token matching system into Dick would understand that each token produced for the query on the Requester's Server matches a corresponding token in Medical Information Repository's de-identified patient

records. In that way the relevant de-identified healthcare information related to patients in a clinical trial can be retrieved in compliance with HIPAA.

xxx. [7.f] the produced healthcare information is absent personally identifiable information and

245. Dick and Landi both describe or suggest that the produced healthcare information is absent personally identifiable information because it is de-identified. Ex. 1007, [0013], [0044]-[0047]; Ex. 1008, 5:32-47.

246. Additionally, Dick and Landi's are both concerned with protecting individual identity in compliance with HIPAA. This would include producing healthcare information without personally identifiable information. Ex. 1007, [0044]-[0047]; Ex. 1008, 5:32-47, 14:26-32.

xxxi. [7.g] is based on a minimum number of matched tokens,

247. This element would have been obvious to a POSA as a further level of protection to comply with HIPAA regulations. It would have been a simple design choice to ensure that a minimum number of tokens are identified before production. If only one token was produced or identified, then there is a significant risk that an individual could be identified based on the selected parameters.

248. As discussed above, a POSA would understand that one way to ensure compliance with 45 C.F.R. § 164.514(b) would be to restrict the query for patient information to at least a minimum number of patients so that information returned could not be associated with a specific individual. 45 C.F.R. § 164.514(b), which

governs the use and disclosure of protected information, provides that “a person with appropriate knowledge and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable” can apply those principles to determine that the “risk is very small that the information could be used, alone or in combination with other reasonably available information” to identify the patient. If a search only returns one individual, or a very small number of individuals, then it would not be HIPAA compliant because the risk would not be very small that the information could be used to identify the patient.

249. Landi also describes or suggests identifying at least a minimum number of corresponding tokens when producing the requested healthcare information because Landi is intended to comply with the provisions of HIPAA. Ex. 1008, 2:5-18; 45 C.F.R. § 164.514(b) (“a person with appropriate knowledge and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable”). Similarly, Dick cites to complying with HIPAA using the statistical de-identification method, which a POSA would understand to include optionally using a minimum number of corresponding tokens. Ex. 1007, [0045] (“Where removing such information is not practical or renders the data useless, the medical information repository may instead release medical information upon condition that a person with appropriate

knowledge and experience with generally accepted statistical and scientific principles has applied a ‘de-identifying’ methodology to remove or make certain that no personally identifiable information is released from the repository.”), [0046].

250. Dick and Landi in combination with Murphy render this limitation obvious. Murphy provides more detail about statistical de-identification and subsequent querying of that data, describes this can be accomplished in compliance with HIPAA to prevent reverse engineering of the patient identity from the produced de-identified healthcare information. Ex. 1009, p. 555 (“One vulnerability is that we allow a large number of queries to be performed that return zero patients. Unfortunately, naïve users who are performing incorrect queries will often have zero patients returned. Having the table-watching agent terminate access to these users would be too sensitive a response. However, attacks around zero are more likely to succeed. Therefore, we modify our range algorithm to have a wider obfuscation at values close to zero, and indicate the greater range by returning ‘<3’ when the obfuscated number is less than 3”). Murphy recognized the potential for misuse when an individual searches using very specific parameters to intentionally return a small number of individuals. To combat this problem, Murphy describes an algorithm designed to prevent the user from determining the identity of the individual patients. One of the safeguards used is to

block any queries that return fewer than 3 patients in order to prevent the user from finding the patient identity. It would have been obvious that Murphy's algorithm could have been combined with Dick and Landi to ensure that a minimum number of tokens are identified because Murphy provides more detail on the same concern about statistical de-identification described in Dick and Landi, and therefore a POSA would be motivated to combine Murphy with Dick and Landi with a reasonable expectation of success.

251. It would have been a simple design choice with predictable results to determine whether the number of individuals returned was at least a minimum number upon receiving the request, upon matching the tokens, or both to ensure compliance with HIPAA.

xxxii. [7.h] the data server is also configured to provide the produced healthcare information to report the characteristics of the group

252. Dick describes a data server, called a Medical Information Repository, for producing the healthcare information that characterizes a group of individuals. Ex. 1007, Fig. 3, [0019], [0028] ("Following receipt of the request for information 105, the Central Server 245 will transmit a query to a medical information repository 275 for information pursuant to the request 115."). In addition, as shown in FIG. 3, an Archive Medical Information System (AMIS) server 265 may be accessed by the central server 245. *Id.*, Fig. 3, [0028]-[0031].

AMIS server 265 will receive the transmitted query 115 and proceed to search the AMIS database for information in satisfaction of the query. *Id.* A POSA would implement Landi's token system and central database of de-identified healthcare data into Dick's system for producing healthcare information. Per Landi, the system operates at different clinical sites and utilizes central patient data processing system to collect de-identified data records. *Id.*, 4:19-28; 5:40-44, 14:26-32. A POSA would have understood that a data server could produce the requested healthcare data because it was a common element in computer systems operating across multiple sites. Additionally, a POSA would have known that the central patient data processing system in Landi, which collects de-identified patient records, would work to provide de-identified healthcare data in response to the query from Dick's Requester's Server.

xxxiii. [7.i] absent authorization from the individuals in the group.

253. Dick and Landi both describe or suggest that the produced healthcare information is absent personally identifiable information because it is de-identified. Ex. 1007, [0013], [0044]-[0047]; Ex. 1008, 5:32-47.

254. Additionally, Dick and Landi's are both concerned with protecting individual identity in compliance with HIPAA. This would include producing healthcare information without personally identifiable information. Ex. 1007, [0044]-[0047]; Ex. 1008, 5:32-47, 14:26-32.

255. Further, as discussed above, Dick seeks to provide access to de-identified healthcare information in a manner that complies with HIPAA, as an alternative to getting explicit patient authorization. Without authorization, it could violate HIPAA to just send PII to be used in querying data records. So, in order to avoid potentially violating HIPAA, a POSA would have used a public encryption key to create tokens on a requesting server. Landi further provides details on how the token creation and matching system works, which a POSA would understand would be an obvious way to implement the de-identification called for by Dick.

xxxiv. [8] The system of claim 7, wherein the provided healthcare information is in conformity with the Health Insurance Portability and Accountability Act (HIPAA).

256. Dick in view of Landi and Murphy describes producing healthcare information absent authorization in accordance with HIPAA. See 7.b above.

257. Further, a POSA would find this limitation is obvious where Dick, Landi, and Murphy are directed to de-identifying health information to comply with HIPAA. See 7.b above.

xxxv. [9] The system of claim 7, wherein the request for healthcare information that characterizes a group of individuals includes personally identifiable information associated with the individuals of the group.

258. In my opinion, Dick in view of Landi and Murphy describes or suggests 9 for the reasons cited in 7.d above.

xxxvi. [10] The system of claim 7, in which the data that identifies each individual is encrypted.

259. In my opinion, Dick in view of Landi and Murphy describes or suggests 10 for the reasons cited in 7.b above.

xxxvii. [11] The system of claim 7, in which the tokens associated with the healthcare data and the request tokens are similarly encrypted.

260. Landi describes or suggests that tokens that are similarly encrypted using the same public encryption key can be used to match the same patient records at different times. Ex. 1008, 8:47-55 (“The public keys are freely distributable (no HIPAA issues) to healthcare organizations for de-identification of patient data. Moreover, the public keys can be used to produce a unique, reproducible encrypted version of patient identifying information. Since the encryption is reproducible, the same patient can be recognized at different times . . . even though the identity of the person remains unknown.”), 4:65-5:2 (“each encryption system (22-1~22-n) implements the same secure public encryption protocol (e.g., RSA).”), 5:32-36 (“The encryption systems (22-1~22-n) can operate to generate Encrypted IDs and then store each Encrypted ID with or within a corresponding de-identified data patient data record”), 5:7-13. Therefore, a POSA implementing Landi’s token scheme into Dick’s system for producing healthcare information for a group of individuals would understand that to match

the produced tokens in the query to the tokens in the patient records in the central data server, they would be similarly encrypted using the same or similar public encryption key. This type of matching token system for querying de-identified healthcare data was well-known at the time and was the primary method used to provide healthcare information between entities without the need to obtain explicit authorization from the patients but still comply with HIPAA.

xxxviii. [12] The system of claim 7, in which the data server provides the requested healthcare information if the group includes at least the minimum number of individuals.

261. In my opinion, Dick in view of Landi and Murphy describes or suggests 12 for the reasons cited in 7.c above.

xxxix. [13] The system of claim 7, in which the healthcare data represents medical related information associated with the individuals of the group.

262. Dick describes that the query from the Requester's Server is for healthcare information that characterizes the group of individuals in the query. Abstract ("The present invention relates to a method of searching for medical information. . . compris[ing] the steps of receiving a request for medical information including identification of a subject; transmitting a query to a medical information repository for information pursuant to the request;"). Dick explains that its system can be used to obtain healthcare information for many purposes, e.g., in connection with clinical trials for new pharmaceuticals, and health

insurance risk assessment for policy underwriting. Ex. 1007, [0003], [0026], [0037], [0041]-[0042], [0047]. Landi also suggests that the healthcare data represents medical related information associated with the individuals of the group because the patient data processed by the system is from healthcare organizations and can be used for various purposes including research public health, and healthcare operations. Ex. 1008, 4:19-23 (“The patient data processing systems (20-1~20-n) are operated at different clinical sites by different healthcare organizations (including, but not limited to, doctors, healthcare providers, institutions, associations, organizations, hospitals, etc.).”), 8:61-9:19 (“Systems and methods according to the invention for deidentifying patient data can be implemented for various purposes such as research, public health or healthcare operations.”).

- xl. [1.a] A computer-implemented method comprising: receiving a request from a user device for information that characterizes a group composed of multiple individuals from a set of de-identified medical data representing characteristics of the individuals pre- or post-encounter with a healthcare facility,**

263. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.a for the reasons cited in 7.a, 7.b, 9 above.

- xli. [1.b] wherein the request message for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,**

264. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.b for the reasons cited in 7.a, 7.b above.

xlii. [1.c] and wherein the request message includes a batch number and at least one report option;

265. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.c for the reasons cited in 7.h above.

266. Dick also suggests that a request message to the data server would include a batch number. For example, Dick explains that each the request can include an identifier for a group of individuals. Ex. 1007, [0024] (“[e]ach request for medical information includes identification of a subject A subject might consist of a . . . group of humans.”). It would be obvious to a POSA that the described subject identifier could be a batch number or any other means of identifying the group. It would be obvious to a POSA to group the requested data using a group identifier or batch number because grouping a requested data was well-known at the time. It would also be obvious to a POSA that the group name or batch number would also be in the request so the user could understand the results. Dick suggests that there may be multiple requesters, multiple encryption servers, and multiple data servers, and multiple requests to the data servers can be made at once. Id. at [0029]-[0030]. Therefore, a POSA would understand that to manage multiple queries and multiple data requests across multiple data servers, one would need to implement an identification scheme in order to maintain and identify which

query a request is related to. As discussed above, it would have been obvious to a POSA to do this including an identification number, such as batch number, in the requests, queries, and responses. Further, Dick describes that “[i]f more than one query was made, central server [] will compile the responses to the queries prior to returning the response to the request.” *Id.* at [0032]. It would have been obvious to a POSA to use batching and a batch number to compile the responses to the request. This is a standard queuing technique that has been used in computer programming since the first programmable computers were invented.

267. Dick also suggests that the request can include one report option in the form of reformatting. *Id.*, [0034] (“Central server [] may provide several additional value added functions to the response(s) to queries prior to transmitting a response to the request []. These value added functions include but are not limited to reformatting, searching for and appending names . . . , and providing interpretive information.”) It would have been obvious to a POSA to implement these value added features using report options. Dick acknowledges that the requests will come from many different reasons for requesting medical information. *Id.* at [0037]. Therefore, a POSA would understand that in order to accommodate the different needs of the different requesters, one would need to allow the requester to specify how the information should be reported (e.g. reformatted, appended, or interpreted) in the form of including a report option in the request.

- xliii. [1.d] producing, by a processor of an encryption server, a unique request token from identifying data for each individual included in the group,**

268. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.d for the reasons cited in 7.b, 7.c, 7.e, and 11 above.

- xliv. [1.e] wherein the request tokens are encrypted before being transmitted over a network;**

269. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.e for the reasons cited in 11, 7.a above.

- xliv. [1.f] comparing the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens;**

270. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.f for the reasons cited in 7.e above.

- xlvi. [1.g] producing the healthcare information with the identity of the individuals being unattainable from the produced healthcare information**

271. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.g for the reasons cited in 7.f above.

- xlvii. [1.h] based on a minimum number of matched tokens; and**

272. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.h for the reasons cited in 7.g above.

- xlvi. [1.i] providing the produced healthcare information to report the characteristics of the group absent authorization from the individuals in the group.**

273. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.i for the reasons cited in 7.h above.

- xlix. **[2] The computer-implemented method of claim 1, wherein the provided healthcare information is in conformity with the Health Insurance Portability and Accountability Act (HIPAA).**

274. In my opinion, Dick in view of Landi and Murphy describes or suggests 2 for the reasons cited in 8 above.

- l. **[3] The computer-implemented method of claim 1, wherein receiving the request for healthcare information that characterizes the group of individuals includes receiving personally identifiable information associated with the individuals of the group.**

275. In my opinion, Dick in view of Landi and Murphy describes or suggests 3 for the reasons cited in 9 above.

- li. **[4] The computer-implemented method of claim 1, in which producing the requested healthcare information includes encrypting respective data that identifies each individual.**

276. In my opinion, Dick in view of Landi and Murphy describes or suggests 4 for the reasons cited in 10 above.

- lii. **[5] The computer-implemented method of claim 1, in which the tokens associated with healthcare data and the request tokens are similarly encrypted.**

277. In my opinion, Dick in view of Landi and Murphy describes or suggests 5 for the reasons cited in 11 above.

- liii. **[6] The computer-implemented method of claim 1, in which the healthcare data represents medical related information associated with the individuals of the group.**

278. In my opinion, Dick in view of Landi and Murphy describes or suggests 6 for the reasons cited in 13 above.

- liv. **[14.a] One or more computer readable storage devices storing instructions that are executable by a processing device, and upon such execution cause the processing device to perform operations comprising:**

279. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.a for the reasons cited in 7.a above.

- lv. **[14.b] receiving a request from a user device for information that characterizes a group composed of multiple individuals from a set of de-identified data representing characteristics of the individuals pre-or post-encounter with a healthcare facility,**

280. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.b for the reasons cited in 7.b above.

- lvi. **[14.c] wherein the request message for healthcare information is sent to an encryption server that is configured to determine if the number of individuals in the request is at least a defined minimum number of individuals,**

281. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.c for the reasons cited in 7.c above.

- lvii. **[14.d] and wherein the request message includes a batch number and at least one report option;**

282. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.d for the reasons cited in 1.c above.

lviii. [14.e] producing, by a processor of an encryption server, a unique request token from identifying data for each individual included in the group,

283. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.e for the reasons cited in 7.d and 11 above.

lix. [14.f] wherein the request tokens are encrypted before being transmitted over a network;

284. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.f for the reasons cited in See 7.a and 11 above.

lx. [14.g] comparing the request tokens to tokens associated with healthcare data that includes de-identified information to find matching tokens;

285. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.g for the reasons cited in 7.e above.

lxi. [14.h] producing the healthcare information with the identity of the individuals being unattainable from the produced healthcare information

286. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.h for the reasons cited in 7.f and 7.g above.

lxii. [14.i] based on a minimum number of matched tokens; and

287. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.i for the reasons cited in 7.g above.

- lxiii. [14.j] providing the produced healthcare information to report the characteristics of the group absent authorization from the individuals in the group.**

288. In my opinion, Dick in view of Landi and Murphy describes or suggests 15 for the reasons cited in 7.h and 7.i above.

- lxiv. [15] The computer readable storage devices of claim 14, wherein the provided information is medical data in conformity with the Health Insurance Portability and Accountability Act (HIPAA).**

289. In my opinion, Dick in view of Landi and Murphy describes or suggests 15 for the reasons cited in 8 above.

- lxv. [16] The computer readable storage devices of claim 14, wherein receiving the request for healthcare information that characterizes the group of individuals includes receiving personally identifiable information associated with the individuals of the group.**

290. In my opinion, Dick in view of Landi and Murphy describes or suggests 16 for the reasons cited in 9 above.

- lxvi. [17] The computer readable storage devices of claim 14, in which producing the requested healthcare information includes encrypting respective data that identifies each individual.**

291. In my opinion, Dick in view of Landi and Murphy describes or suggests 17 for the reasons cited in 10 above.

- lxvii. [18] The computer readable storage devices of claim 14, in which the tokens associated with the healthcare data and the request tokens are similarly encrypted.**

292. In my opinion, Dick in view of Landi and Murphy describes or suggests 18 for the reasons cited in 11 above.

lxviii. [19] The computer readable storage devices of claim 14, in which the healthcare data represents medical related information associated with the individuals of the group.

293. In my opinion, Dick in view of Landi and Murphy describes or suggests 19 for the reasons cited in 13 above.

lxix. [20] The computer readable storage devices of claim 14, in which producing the requested healthcare information includes determining if the group includes at least the minimum number of individuals.

294. In my opinion, Dick in view of Landi and Murphy describes or suggests 19 for the reasons cited in 7.c above.

lxx. [21] The computer-implemented method of claim 1, wherein the identity of the individuals is unattainable from the unique request tokens.

295. In my opinion, Dick in view of Landi and Murphy describes or suggests 20 for the reasons cited in 7.f and 7.g above.

lxxi. [22] The computer-implemented method of claim 1, wherein the unique request tokens and the tokens associated with the healthcare data are similarly encrypted.

296. In my opinion, Dick in view of Landi and Murphy describes or suggests 20 for the reasons cited in 11 above.

X. CONCLUSION

- h. It is my opinion that a POSA would have found the de-identification of sensitive data, the subject matter of claims 1-20 of the '375 Patent to be both well-known and obvious at the time of the purported invention.
- i. I have reached the foregoing opinions based on my knowledge, education, experience, and expertise, which enables me to speak to what a POSA would have understood at the relevant time period.
- j. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.



Paul C. Clark DSc.

Dec. 14, 2023