UNITED STATES PATENT AND TRADEMARK OFFICE

_____

BEFORE THE PATENT TRIAL AND APPEAL BOARD

_____

DATAVANT, INC.

Petitioner

v.

VIGILYTICS LLC

Patent Owner

IPR2024-00382

Patent No. 9,665,685

Issue Date: May 30, 2017

Title: Using De-Identified Healthcare Data to Evaluate Post-Healthcare Facility Encounter Treatment Outcomes

---

DECLARATION OF PAUL C. CLARK, DSC. IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 9,665,685

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 ’685 Patent 15
- A. Background 15

IX. The Grounds of Unpatentability 18
- A. Prior Art References 18
- B. Ground 1: Obviousness (§103) over Evenhaim 36
- C. Ground 2: Obvious (§ 103) over Evenhaim in view of Settimi 56
- D. Ground 3: Obvious (§ 103) over Dick in view of Landi and Murphy 74

X. Conclusion 107

Datavant Exhibit 1003

EXHIBIT LIST

Exhibit Number	Description
1001	U.S. Patent No. 9,665,685 (the "'685 Patent")
1002	File History of U.S. Patent No. 9,665,685
1003	Declaration of Dr. Paul Clark ("Clark")
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, *A Security Architecture for Query Tools used to Access Large Biomedical Databases*, 2002 AMIA Annual Symposium Proceedings, 552 (2002) ("Murphy")
1010	*Milliman, Inc. v. Gradient A.I. Corp.*, No. 1:21-Cv-10865-NMG (D. Mass.), Dkt. 108.
1011	*Credit Card Security: The Best Way to Secure Data is Not to Store the Data* (Hospitality Upgrade, Summer 2008, www.hospitalityupgrade.com)
1012	Jennifer G. Steiner, B. Clifford Neuman, and Jeffrey I. Schiller, *Kerberos: An Authentication Service for Open Network Systems*, 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. 9,665,685 (the "'685 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. 9,665,685 (the "'685 Patent")
1002	File History of U.S. Patent No. 9,665,685
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, *A Security Architecture for Query Tools used to Access Large Biomedical Databases*, 2002 AMIA Annual Symposium Proceedings, 552 (2002) ("Murphy")
1010	*Milliman, Inc. v. Gradient A.I. Corp.*, No. 1:21-Cv-10865-NMG (D. Mass.), Dkt. 108.
1011	*Credit Card Security: The Best Way to Secure Data is Not to Store the Data* (Hospitality Upgrade, Summer 2008, www.hospitalityupgrade.com)

1012	Jennifer G. Steiner, B. Clifford Neuman, and Jeffrey I. Schiller, *Kerberos: An Authentication Service for Open Network Systems*, 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 ’685 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 ’685 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 ’685 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 '685 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 "data" is used throughout the Challenged Claims. Based on my reading of the '685 Patent a POSA would understand that "data" as it is used throughout the patent to mean "medical information associated with the individuals of the group."

28. For example, claim 13 says that "medical 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 1:41-43. 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
- Medical information . . . – 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 '685 Patent. Ex. 1001, 2:6-8 ("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 1-20 of the '685 Patent are obvious over Evenhaim.

32. Claims 1-20 of the '685 Patent are obvious over Evenhaim in view of Settimi.

33. Claims 1-20 of the '685 Patent are obvious over Dick in view of Landi and Murphy.

VIII. THE ’685 PATENT

A. Background

34. The ’685 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, 4:39-44. 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, 4:45-5:2. For example, healthcare facilities need to get the healthcare information for a group of individuals to evaluate the risk of rehospitalization, but getting explicit authorization from every individual 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 '685 patent is to provide a method for complying with HIPAA using a token-based system for de-identifying patient

records. The limitations in the ’685 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 ’685 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 ’685 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 relates to data privacy and usage in a distributed system 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, 1:14-19, 5:4-7.

39. Evenhaim's system includes a secure aggregator, data sources (e.g. healthcare facilities), and source entity 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 users 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 allows 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 known 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 the system shown in Fig. 4 for querying healthcare information, the specification describes that encryption can take place at the data aggregator in Fig. 4 as well as at 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 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 source 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 '685 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 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 at the data aggregator, 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.

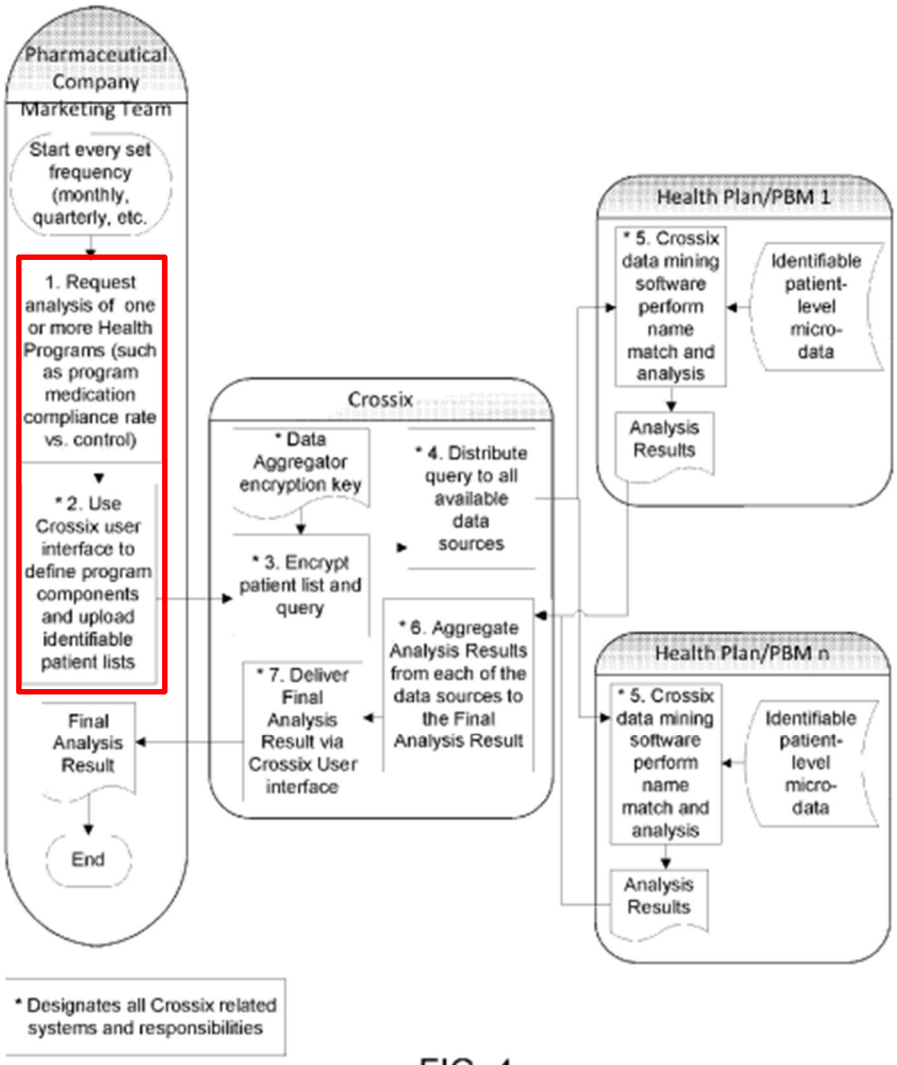

FIG. 4

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.

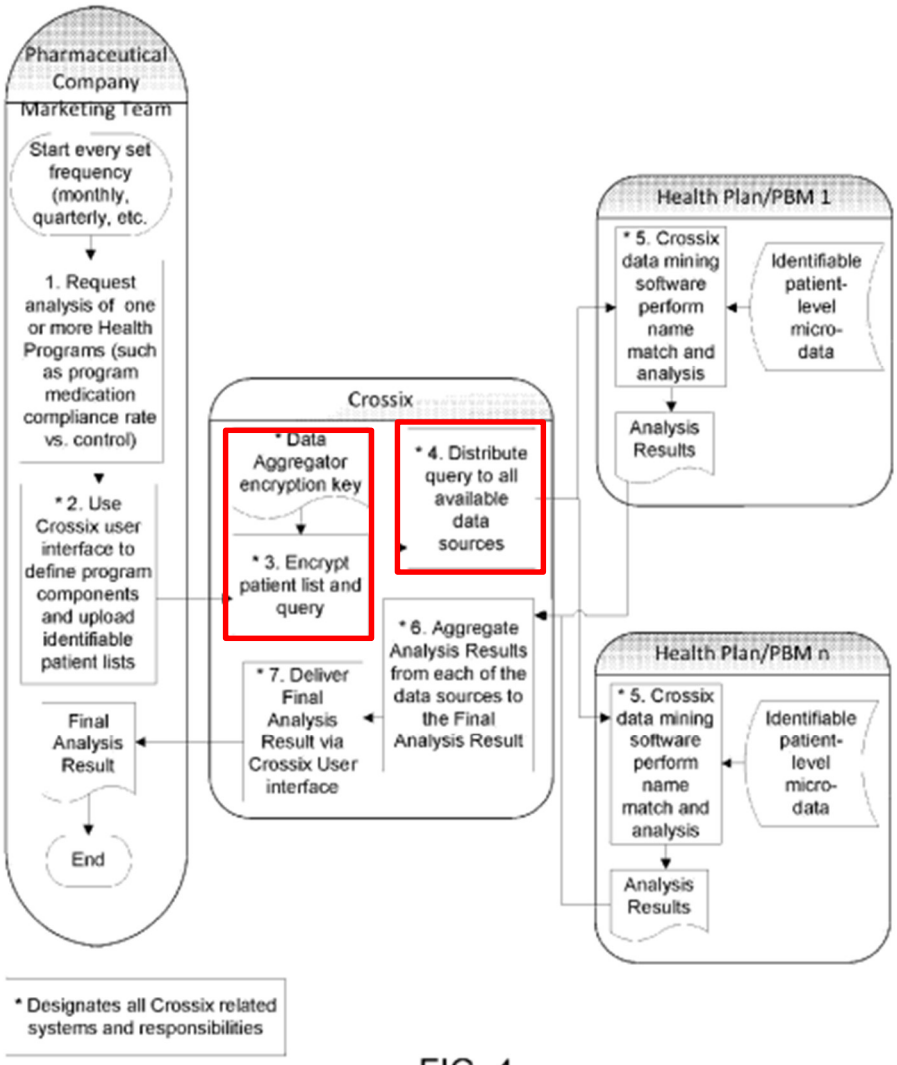

FIG. 4

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. Evenhaim's

system is also flexible such that encryption can happen on either the aggregator processor or the source-entity processor. 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.

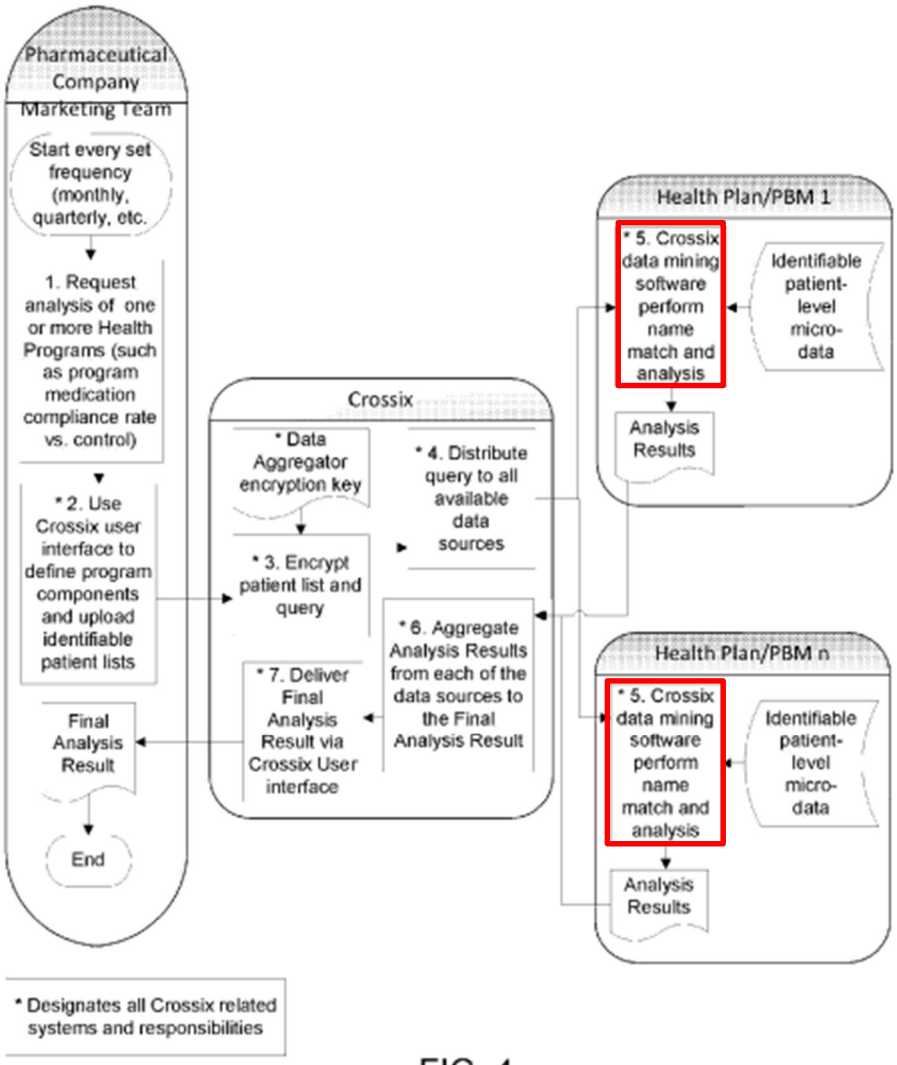

FIG. 4

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 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 '685 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 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 from a client device 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, 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 '685 patent.

57. Landi was not considered by the Patent Office during prosecution of the '685 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:

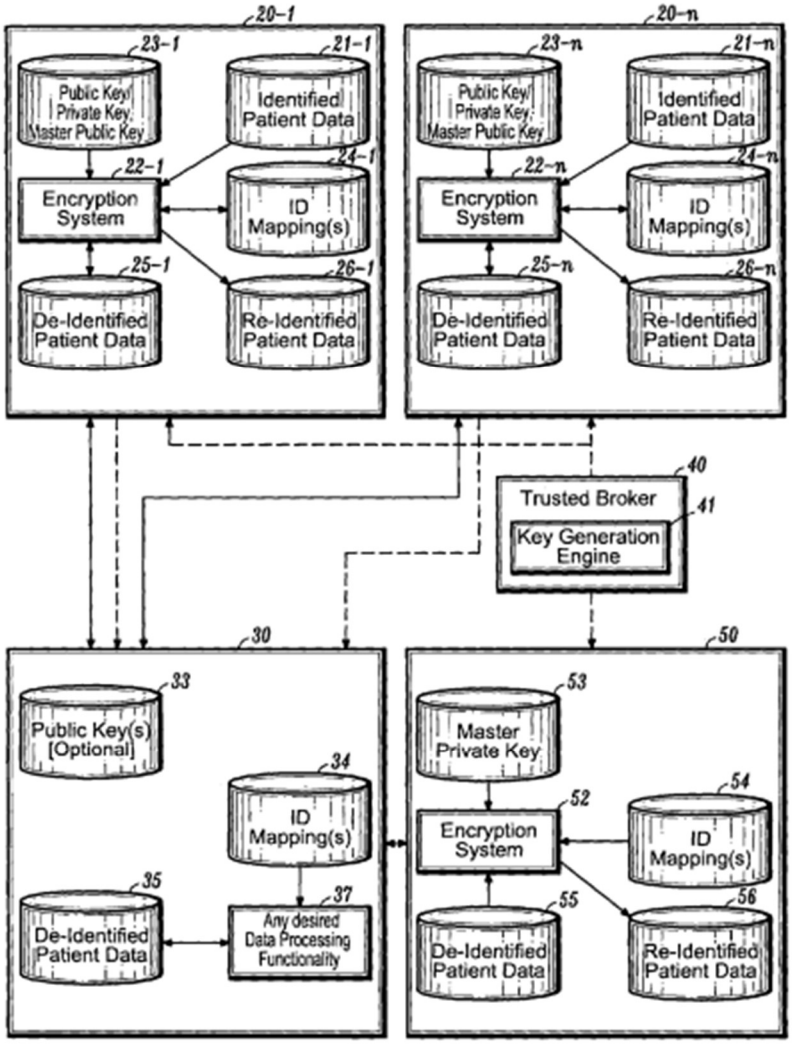

FIG. 1

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: Obviousness (§103) over Evenhaim

67. In my opinion, each element of the '685 patent's claims 1-20 would be obvious to a POSA over Evenhaim. Support for my opinions is detailed further below.

i. [7.a] A system comprising: a computing device comprising: a memory configured to store instructions; and a processor to execute the instructions to perform operations comprising:

68. Evenhaim's system is implemented on computers (i.e. computing devices). Ex. 1005, 14:64-66. Evenhaim's system also includes a server that sends the user's request for information called a "data aggregator," as well as a data server that responds to the request for information called "data source," "source entity" or "data source entity," both of which connect with other machines using, e.g., the internet. Ex. 1005, 6:34-42 ("interconnected with each other using standard data-communications facilities (e.g. Internet, VPN, etc.)."); *see also* 10:64-11:2, 17:53-59.

69. Both the data aggregator and source-entity have processors for executing a protocol. *Id.*, 6:57-63; 17:53-63.

70. A POSA would have found it obvious that the computing device would have a memory because (i) computing devices routinely use memories to

store instruction for performing functions and (ii) Evenhaim describes software for executing the instructions, embedded in memory media, to perform operations. *Id.*, 17:53-64 ("memory media having any of said software encoded therein"). 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 with a memory. 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.

71. Therefore, it would be obvious that the computing device would have memory because Evenhaim describes that software for executing the protocols is encoded in memory media. *Id.*, at 17:53-64.

ii. [7.b] receiving a request from a user device for medical information that characterizes a group composed of multiple individuals

72. Evenhaim describes that its data aggregator can receive requests from a user interface for various types of health information that characterizes a group composed of multiple individuals. Ex. 1005, 7:65-8:8:4, 8:65-9:12; 15:3-14; 15:25-45, 8:49-64. The request can be for information that characterizes a group of multiple individuals. *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."); *see also* Figs. 4, 4A, 5, 7:65-8:4, 8:18-22, 8:30-41, 8:65-9:12, 15:3-14, 15:25-45.

73. Examples include identifying common treating physicians for the group, tracking the health outcomes of patients in a clinical trial, etc. *Id.*, 8:65-9:12; 15:3-14; 15:25-45.

iii. [7.c] from a set of de-identified medical data representing pre- or post-encounter characteristics of the individuals

74. Evenhaim describes that its system can be used to request characteristics of a group of individuals that relate to encounters with healthcare facilities to assess the impact of health programs. *Id.*, 15:3-9, Fig. 4A. 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.

75. The request is for healthcare information that characterizes the group, including the de-identified patient records themselves, identifying common treating physicians, and tracking outcomes of patients in a clinical trial. Ex. 1005, Fig. 5, 7:12-26, 8:65-9:12, 15:3-14, 15:25-45.

76. A POSA would understand the de-identified data is information about a group of individuals after (post) an encounter with a healthcare facility because this data includes diagnosis, treatment, pharmacy, and clinical data. *Id.*, 15:3-20,

Fig. 4A. A pharmacy is a well-known healthcare facility. Additionally, healthcare facilities are responsible for diagnosing, treating, and compiling clinical data. Evenhaim also describes comparing the information from individuals who took part in a health program to a control group. *Id.*, 15:9-20, Fig. 4A. A POSA would understand that a control group may include individuals before they encountered a healthcare facility and entered a health program. A POSA would also understand that the data is de-identified because these healthcare facilities are under strict privacy standards like HIPAA.

iv. [7.d] wherein the request contains identifying data for each member in the group of multiple individuals;

77. Evenhaim describes that the request from the user interface may include a "parameter list," which may include individually identifying information such as names, dates of birth, etc. Ex. 1005, 13:62-14:13, 8:39-43 ("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. See also my arguments with respect to element 7.c.

v. [7.e] producing a request token for each individual included in the group of individuals;

78. Evenhaim describes utilizing an encryption key for creating a token at the data aggregator processor. 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 understanding 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; 17:53-63. Therefore, Evenhaim suggests creating a request token for each individual included in the group of individuals. The token is created by encrypting the identifying information using an encryption key. *Id.*, 12:4-16, 12:38-45, 13:4-12.

79. 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.

80. 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 ’685 patent.

81. Thus, Evenhaim describes replacing personal identity information in a population, requesting searches regarding properties of a group of individuals (requesting a query), and producing results for 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 ’685 patent.

vi. [7.f] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

82. Evenhaim describes or suggests tokens in a set of de-identified healthcare data and that the associated de-identified healthcare data corresponds to the produced tokens. A POSA would understand that the tokens created for the query in Figs. 4 and 4A on the data aggregator (the produced tokens) are sent to the data sources, where they are matched against tokens in the patient records in order access the associated de-identified healthcare data that corresponds to those patients (see boxes 3-5). *Id.*, 12:13-16, 12:4-47 (explaining that the encryption key is used on the data source side to de-identify the records). For example, in Evenhaim a user requests Health Program information such as medication compliance rate, the system uses a data aggregator encryption key to encrypt the

patient list and query, distributes it to all data sources, and then performs a match and analysis on patient data before delivering results. Evenhaim further describes that the healthcare data can be accessed using tokens, utilizing a unique encryption key. *Id.*, 6:57-7:26, 8:30-38, 9:1-2, 9:47-51, 12:38-45, 16:5-21.

83. As described above, a POSA would understand that the token creation and matching system described in Evenhaim, *Id.*, 12:4-13:20, could be applied to the embodiment in Figs. 4 and 4A such that the data sources in box 6 would have de-identified patient records with healthcare information with tokens that match the tokens sent with the query in box 4. The specification describes that the token creation could be performed on the aggregator as well as on the data sources. *Id.*, 12:26-34, 17:53-63. As noted above, Evenhaim discloses some embodiments for producing healthcare information that do not require matching tokens. Rather, in some embodiments, the de-identification and creation of tokens takes place after the records are matched. Ex. 1005, Fig. 5, 17:42-52. However, a POSA would understand that for embodiments that the relevant records could be located by matching tokens, which is even acknowledged by Evenhaim. Ex. 1005, 12:4-16; 12:27-33 ("[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 … or on the 'source-entity' processors' side …. "). Therefore, a POSA applying the features of the embodiment of Fig. 4 would have

produced tokens created from the parameter list received by the data aggregator, using the common encryption key, to be used in the query. Otherwise, the operation shown in Fig. 4 would require repeated encrypting and decrypting of the patient names. To comply with HIPAA, the embodiment would encrypt the patient names on the data aggregator side, send the query to the data source, decrypt the patient names in order to perform name matching, and then de-identify the located records using encryption in order to send the results back to the data aggregator side, a more cumbersome and less safe process.

vii. [7.g] producing the requested medical information

84. The disclosures cited for 7.b and 7.d also show that 7.g is obvious. Evenhaim describes providing produced healthcare information to report the characteristics of the group, in the form of reporting the medical information for the group. Ex. 1005, 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

viii. [7.h] with the identity of the individuals being unattainable from the produced medical information;

85. 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. 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.

ix. [7.i] without the individuals in the group having authorized production of such medical information that characterizes the group;

86. 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. [7.j] and providing the produced medical information to report the characteristics of the group of individuals

87. The disclosures for elements 7.b and 7.e also show this element is obvious. Evenhaim describes providing produced healthcare information to report the characteristics of the group, in the form of reporting the medical information

for the group. Ex. 1005, 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

xi. [7.k] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

88. Evenhaim's system describes or suggests producing healthcare information without individuals in the group having authorized its release. As the'641 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.

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

89. The disclosures cited for 7.k also show that claim 8 is obvious.

xiii. [9] The system of claim 7, wherein the request for medical information that characterizes a group of individuals includes

personally identifiable information associated with the individuals of the group.

90. 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.

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

91. 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.

92. 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.

93. 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.

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

94. 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 . . . .").

95. 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.

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

96. 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. In view of Evenhaim and the requirements of HIPAA, a POSA would have also found it obvious to only provide the requested medical information if the group includes a minimum number of individuals. A POSA would have understood that ensuring 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.

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

97. 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.

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

98. In my opinion, Evenhaim describes or suggest 1.a for the reasons cited in 7.a, 7.b, 7.c above.

xix. [1.b] wherein the request contains identifying data for each member in the group of multiple individuals;

99. In my opinion, Evenhaim describes or suggests 1.b for the reasons cited in 7.d above.

xx. [1.c] producing, by a hardware processor, a request token for each individual included in the group of individuals;

100. In my opinion, Evenhaim describes or suggests 1.c for the reasons cited in 7.e above.

xxi. [1.d] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

101. In my opinion, Evenhaim describes or suggests 1.d for the reasons cited in 7.f above.

xxii. [1.e] producing the requested medical information,

102. In my opinion, Evenhaim describes or suggests 1.e for the reasons cited in 7.g above.

xxiii. [1.f] the identity of the individuals being unattainable from the produced medical information

103. In my opinion, Evenhaim describes or suggests 1.f for the reasons cited in 7.h above.

xxiv. [1.g] without the individuals in the group having authorized production of Such medical information that characterizes the group; and

104. In my opinion, Evenhaim describes or suggests 1.g for the reasons cited in 7.i above.

xxv. [1.h] providing the produced medical information to report the characteristics of the group of individuals

105. In my opinion, Evenhaim describes or suggests 1.h for the reasons cited in 7.j above.

xxvi. [1.i] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

106. In my opinion, Evenhaim describes or suggests 1.j for the reasons cited in 7.k above.

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

107. In my opinion, Evenhaim describes or suggests 2 for the reasons cited in 8 above.

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

108. In my opinion, Evenhaim describes or suggests 3 for the reasons cited in 9 above.

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

109. In my opinion, Evenhaim describes or suggests 4 for the reasons cited in 10 above.

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

110. In my opinion, Evenhaim describes or suggests 5 for the reasons cited in 11 above.

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

111. In my opinion, Evenhaim describes or suggests 6 for the reasons cited in 13 above.

xxxii. [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:

112. In my opinion, Evenhaim describes or suggests 14.a for the reasons cited in 7.a above.

xxxiii. [14.b] receiving a request from a user device for medical information that characterizes a group composed of multiple individuals from a set of de-identified medical data

representing pre- or post-encounter characteristics of the individuals,

113. In my opinion, Evenhaim describes or suggests 14.b for the reasons cited in 7.b and 7.c above.

xxxiv. [14.c] wherein the request contains identifying data for each member in the group of multiple individuals who have had an encounter with a healthcare facility;

114. In my opinion, Evenhaim describes or suggests 14.c for the reasons cited in 7.a and 7.d above.

xxxv. [14.d] producing, by a hardware processor, a request token for each individual included in the group of individuals;

115. In my opinion, Evenhaim describes or suggests 14.d for the reasons cited in 7.e above.

xxxvi. [14.e] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

116. In my opinion, Evenhaim describes or suggests 14.e for the reasons cited in 7.f above.

xxxvii. [14.f] producing the requested information

117. In my opinion, Evenhaim describe or suggest 14.f for the reasons cited in 7.g above.

xxxviii. [14.g] with the identity of the individuals being unattainable from the produced medical information

118. In my opinion, Evenhaim describe or suggest 14.g for the reasons cited in 7.h above.

xxxix. [14.h] without the individuals in the group having authorized production of such medical information that characterizes the group; and

119. In my opinion, Evenhaim describe or suggest 14.h for the reasons cited in 7.i above.

xl. [14.i] providing the produced medical information to report the characteristics of the group of individuals

120. In my opinion, Evenhaim describe or suggest 14.i for the reasons cited in 7.j above.

xli. [14.j] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

121. In my opinion, Evenhaim describes or suggests 14.j for the reasons cited in 7.k above.

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

122. In my opinion, Evenhaim describes or suggests 15 for the reasons cited in 8 above.

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

123. In my opinion, Evenhaim describes or suggests 16 for the reasons cited in 9 above.

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

124. In my opinion, Evenhaim describes or suggests 17 for the reasons cited in 10 above.

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

125. In my opinion, Evenhaim describes or suggests 18 for the reasons cited in 11 above.

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

126. In my opinion, Evenhaim describes or suggests 19 for the reasons cited in 13 above.

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

127. In my opinion, Evenhaim describes or suggests 20 for the reasons cited in 12 above.

128. As discussed above Petitioner believes that all of the claimed elements are rendered obvious by Evenhaim.

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

129. In my opinion, each element of the ’685 patent’s claims 1-20 is obvious to a POSA over Evenhaim in view of Settimi. Support for my opinions is detailed further below.

130. A POSA would look to combine these three references because they both seek to solve the same problem of providing access to patient data in a manner that is compliant with privacy laws such, as HIPAA, as explained in sections IX.A.i, IX.A.ii, and IX.A.v. 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.

131. 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].

132. 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. These two references 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. 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.

i. [7.a] A system comprising: a computing device comprising: a memory configured to store instructions; and a processor to execute the instructions to perform operations comprising:

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

ii. [7.b] receiving a request from a user device for medical information that characterizes a group composed of multiple individuals

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

iii. [7.c] from a set of de-identified medical data representing pre- or post-encounter characteristics of the individuals,

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

136. Settimi also describes that users may request information that characterizes a group of individuals from a set of de-identified data representing characteristics of the individuals pre-or post-encounter with a healthcare facility. Ex. 1006, [0027] ("Certain embodiments provide a secure process for sending de-identified patient information from an ambulatory patient care provider (PCP) site to a data warehouse system where the patient data may be analyzed and compared with a wider range of patient data. . . . Information obtained through the data warehouse that pertains to individual patients is transmitted back to the originating PCP site, via a cohort report."). A POSA would understand that such patient data is usually recorded after a patient's medical visit with a doctor or other healthcare provider.

iv. [7.d] wherein the request contains identifying data for each member in the group of multiple individuals;

137. Per my discussion for element 7.d in Ground 1, Evenhaim describes and/or suggest this element at the relevant time period. Additionally, a POSA would understand that the data includes diagnosis, treatment, pharmacy, and clinical data and this naturally would be after an encounter with a healthcare facility. A POSA would also understand that a control group often includes individuals who have not encountered a healthcare facility.

v. [7.e] producing a request token for each individual included in the group of individuals;

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

139. 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.

140. 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.*

141. 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].

142. 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.

vi. [7.f] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens

143. Per my discussion for element 7.f in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

144. As shown in Figs. 4 and 4A, the produced tokens in the query are sent to the data source (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].

145. When the data source 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 source 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, Settimi 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.

146. A POSA would understand that the token creation and matching in Evenhaim, when combined with the central data warehouse of Settimi, 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.

vii. [7.g] producing the requested medical information,

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

viii. [7.h] with the identity of the individuals being unattainable from the produced medical information;

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

149. 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].

ix. [7.i] without the individuals in the group having authorized production of such medical information that characterizes the group;

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

151. Both Evenhaim and Settimi describe how providing the de-identified healthcare information complies with HIPAA, as an alternative to obtaining explicit patient authorizations from the individuals involved. The purpose of de-identification in accordance with HIPAA is to create shareable data that does not require permission from each patient whose information is shared. 45 C.F.R. § 164.502(d); Ex. 1001, 1:33-36, 4:1-3.

152. Evenhaim suggests that healthcare information can be produced without authorization from individuals in the group in accordance with HIPAA. Ex. 1005, 3:17-31 ("Health transaction data is protected by privacy standards such

as the HIPAA . . . by properly de-identifying transaction data from patient identifiers."), 4:34-48, 5:46-6:10, 8:18-22, 14:7-13 ("In each 'query relationship' there are different legal standards which may be applicable—such as … HIPAA … de-identified without micro-data specific content or substantially equivalently provably statistically intractable."), 15:15-24, 14:59-67.

153. Settimi also suggests that the healthcare information can be produced under HIPAA as an alternative to obtaining explicit authorization. HIPAA's Privacy Rule. Ex. 1006, [0009] ("This removal of identifying information must be performed per the federal Health Insurance Portability and Accountability Act (HIPAA) regulations."), [0014], [0017], [0027] ("Certain embodiments provide a secure process for sending de-identified patient information from a [data source] to a data warehouse system where the patient data may be analyzed . . . . The terms 'de-identified patient information' and 'de-identified patient data' . . . refer to both fully de-identified data as defined by HIPAA and limited data set data as defined by HIPAA."), [0039]-[0040].

x. [7.j] and providing the produced medical information to report the characteristics of the group of individuals

154. Per my discussion for element 7.j in Ground 1, Evenhaim describes or suggests this element at the relevant time period. Additionally, Settimi describes or suggests this limitation as well. Ex. 1006, [0027], [0030], [0035]-[0037], [0039]-[0042].

xi. [7.k] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

155. Per my discussion for element 7.k in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

156. Both Evenhaim and Settimi describe how providing the de-identified healthcare information complies with HIPAA, as an alternative to obtaining explicit patient authorizations from the individuals involved. The purpose of de-identification in accordance with HIPAA is to create shareable data that does not require permission from each patient whose information is shared. 45 C.F.R. § 164.502(d); Ex. 1001, 1:51-54, 4:39-44.

157. Evenhaim suggests that healthcare information can be produced without authorization from individuals in the group in accordance with HIPAA. Ex. 1005, 3:17-31 ("Health transaction data is protected by privacy standards such as the HIPAA . . . by properly de-identifying transaction data from patient identifiers."), 4:34-48, 5:46-6:10, 8:18-22, 14:7-13 ("In each 'query relationship' there are different legal standards which may be applicable—such as … HIPAA … de-identified without micro-data specific content or substantially equivalently provably statistically intractable."), 15:15-24, 14:59-67.

158. Settimi also suggests that the healthcare information can be produced under HIPAA as an alternative to obtaining explicit authorization. HIPAA's

Privacy Rule. Ex. 1006, [0009] ("This removal of identifying information must be performed per the federal Health Insurance Portability and Accountability Act (HIPAA) regulations."), [0014], [0017], [0027] ("Certain embodiments provide a secure process for sending de-identified patient information from a [data source] to a data warehouse system where the patient data may be analyzed . . . . The terms 'de-identified patient information' and 'de-identified patient data' . . . refer to both fully de-identified data as defined by HIPAA and limited data set data as defined by HIPAA."), [0039]-[0040].

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

159. Per my discussion for element 8 in Ground 1, Evenhaim describes or suggests this element at the relevant time period. The disclosures cited for 7.k in this Ground also show that 8 is obvious.

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

160. Per my discussion for claim 9 in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

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

161. Per my discussion for claim 10 in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

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

162. Per my discussion for claim 11 in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

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

163. Per my discussion for claim 12 in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

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

164. Per my discussion for claim 13 in Ground 1, Evenhaim describes or suggests this element at the relevant time period.

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

165. In my opinion, Evenhaim in view of Settimi describes or suggests 1.a for the reasons cited in 7.a, 7.b, 7.c, 7.d above.

xix. [1.b] wherein the request contains identifying data for each member in the group of multiple individuals;

166. In my opinion, Evenhaim in view of Settimi describes or suggests 1.b for the reasons cited in 7.d above.

xx. [1.c] producing, by a hardware processor, a request token for each individual included in the group of individuals;

167. In my opinion, Evenhaim in view of Settimi describes or suggests 1.c for the reasons cited in 7.e above.

xxi. [1.d] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens and

168. In my opinion, Evenhaim in view of Settimi describes or suggests 1.d for the reasons cited in 7.f above.

xxii. [1.e] producing the requested medical information,

169. In my opinion, Evenhaim in view of Settimi describes or suggests 1.f for the reasons cited in 7.fi above.

xxiii. [1.f] the identity of the individuals being unattainable from the produced medical information

170. In my opinion, Evenhaim in view of Settimi describes or suggests 1.g for the reasons cited in 7.j above.

xxiv. [1.g] without the individuals in the group having authorized production of Such medical information that characterizes the group; and

171. In my opinion, Evenhaim in view of Settimi describes or suggests 1.g for the reasons cited in 7.j above.

xxv. [1.h] providing the produced medical information to report the characteristics of the group of individuals

172. In my opinion, Evenhaim in view of Settimi describes or suggests 1.g for the reasons cited in 7.j above.

xxvi. [1.i] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

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

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

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

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

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

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

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

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

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

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

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

xxxii. [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:

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

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

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

xxxiv. [14.c] wherein the request contains identifying data for each member in the group of multiple individuals;

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

xxxv. [14.d] producing, by a hardware processor, a request token for each individual included in the group of individuals;

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

xxxvi. [14.e] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

183. In my opinion, Evenhaim in view of Settimi describes or suggests 14.e for the reasons cited in 7.f above.

xxxvii. [14.f] producing the requested information

xxxviii. [14.g] with the identity of the individuals being unattainable from the produced medical information; and

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

xxxix. [14.h] without the individuals in the group having authorized production of Such medical information that characterizes the group; and

xl. [14.i] providing the produced medical information to report the characteristics of the group of individuals

xli. [14.j] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

185. In my opinion, Evenhaim in view of Settimi describes or suggests 14.i for the reasons cited in 7.k and 7.l above.

xlii. [15] The computer readable storage devices of claim 14, wherein the provided medical information is in conformity

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

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

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

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

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

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

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

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

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

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

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

191. In my opinion, Evenhaim in view of Settimi describes or suggests 20 for the reasons cited in 12 above.

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

192. In my opinion, each element of the '685 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.

193. 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].

194. 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.

195. 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.

196. 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 ground 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 the 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.

197. It would be obvious to a POSA that to use the public key from Landi to create a 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.

198. 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].

199. 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 created 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, 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.

200. A POSA would 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, as explained in Sections IX.A.iii, IX.A.iv, and IX.A.v.

201. 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.

202. 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.

i. [7.a] A system comprising: a computing device comprising: a memory configured to store instructions; and a processor to execute the instructions to perform operations comprising:

203. Dick describes a system that includes a "Requester's Server" (230) that receives a user request for the names of a group of individuals from a "Requesting & Viewing Client," and then creates a query that it transmits, in connection with central server 245, to a "Medical Information Repository." Ex. 1007, Fig. 3, [0020]-[0022], [0028], [0034], claim 23. The query is for de-identified healthcare information in order to comply with HIPAA. Ex. 1006, [0044]-[0046]. Dick's System includes a client-side hardware and software environment that allows for network access to internal and external server-based application. *Id.*, [0020]-[0021] ("Each front-end physical environment includes one or more requesting and viewing clients ('RVC's), respectively, client-server-based 215, intranet-based 220, and internet-based 225."), [0036], [0048] ("general purpose computer", "appropriate hardware", "memory"), [0049]. A POSA would have found it obvious that Dick's Requester's Server has a processor and a memory.

204. Landi describes using encryption to create tokens from PII using a public encryption key. Ex. 1008, 2:63-67, 3:3-5, 8:38-55 ("public keys are freely distributable"), 14:29-32. Landi also establishes that these computerized systems use "processors" and "program storage devices." Ex. 1008, 3:58-4:10.

ii. [7.b] receiving a request from a user device for medical information that characterizes a group composed of multiple individuals

205. Dick's Requester's Server receives a request from the "requesting and viewing client" for 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.

206. 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.

207. 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.*

208. Dick also describes that the query 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 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 key system into Dick's system to create the produced tokens from the personal identifying information specified in the user request. *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 source.

209. Additionally, a POSA would be motivated to implement Landi's key system into Dick's request from the client device because it would be preferred that the request sent to the data source 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.

210. 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.

iii. [7.c] from a set of de-identified medical data representing pre- or post-encounter characteristics of the individuals,

211. Dick explains that the repository responds to queries for information from a set of de-identified data representing characteristics of individuals. *Id.*, [0045] ("Before information from a medical information repository could be released . . . the medical information would have to be altered such that no person can be individually identified from the health information."), 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 upon verification that the request is authorized.").

212. Dick also suggests that the queried information relates to pre or post encounters with health facilities. *Id.*, [0018] ("Medical Information 100 encompasses all information relating to physical and mental health diagnoses and remedies such as physician-patient records, clinical information records and prescription drug records. Clinical information includes laboratory testing, ambulatory, home health, and long-term care among other sources of clinical care and information."); *see also id.*, [0046] ("The information distributed by the present invention can be deidentified at one of any number of points, including

being deidentified at the originating Source of the information, Such as the PBM, or the clinic from which the medical information is being collected."). See also my arguments with respect to element 7.b.

iv. [7.d] wherein the request contains identifying data for each member in the group of multiple individuals;

213. Dick explains that the request for healthcare information identifies individuals in a group for which the data is sought. Ex. 1007, Fig. 1, box 105 ("Receiving a request for medical information, the request including identification of a subject."), [0024] ("[S]ubject[s] might consist of a human individual or group of humans."), [0019]-[0022], claim 23 ("[S]earching at least one medical information repository for medical information regarding a group of preselected individuals."). My analysis for element 7.c also show why this element is obvious.

214. In Landi the request for healthcare information includes PII associated with members of the group because the Encrypted ID that are used in the query are created from PII associated with the patients. Ex. 1008, 4:40-47 ("[T]he public keys … are used for de-identifying (encrypting) one or more items of patient identifying information … to generate an encrypted version of the patient identifying information (referred to herein as 'Encrypted IDs)'."). For example, Landi describes how its invention can be used to monitor patients in a clinical trial for compliance while still complying with HIPAA. Ex. 1008, 9:43-10:15. A POSA would have appreciated that when patients are queried under this scenario, the

tokens are created using the PII for each patient in the trial (i.e. the group in question). Ex. 1008, 9:52-56 ("[E]ach time data is collected, the patient should be de-identified in the same way (e.g., it is critical to know that de-identified patient 24 from a first round of collection is also present in a second round and is still de-identified as de-identified patient 24)."). My arguments for element 7.c also show why this element is obvious.

v. [7.e] producing a request token for each individual included in the group of individuals;

215. 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.

216. 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.

217. 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.*

218. Dick also describes that the request 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 key system 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 key system into Dick's system to create the produced tokens from the personal identifying information specified in the user request. *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 source.

219. Additionally, a POSA would be motivated to implement Landi's key system into Dick's system because it would be preferred that the user request sent to the data source 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 for a request. 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.

220. 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

vi. [7.f] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

221. The tokens produced by Dick's server in response to the user request in queries sent from the Requester's Server in Dick would correspond to tokens inserted into the de-identified patient records in the Medical Information Repository because a POSA would be motivated to use the same or similar public encryption key in accordance with the teaching in Landi. Landi describes tokens in a set of de-identified healthcare data and the associated healthcare data that would correspond to these tokens. Ex. 1008, FIG. 3. Each time patient records are updated at their source, corresponding tokens are inserted into the patient data to replace the personal identifying information, and the de-identified patient data that is sent to the central repository for future querying. *Id.*, 5:32-47, 8:47-55. Therefore, a POSA would understand that in order to query the central database,

the produced tokens should be encrypted using the same or similar public encryption key to the corresponding tokens in patient records in order to be able to match the patient records to the same tokens produced by Dick's server in response to a user request. *Id.*

222. Therefore, a POSA would understand that the token created in response to a user request and sent to the medical repository would be compared against the tokens in the de-identified patient data to find matching patient records.

vii. [7.g] producing the requested medical information

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

viii. [7.h] with the identity of the individuals being unattainable from the produced medical information

224. The identity of the individuals that correspond to the tokens (Encrypted IDs) described in Landi, and implemented on Dick's system, 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).

225. 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 server will produce encrypted tokens from the identifying information provided in the user's request, 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 key system into Dick's system 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.").

226. 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. *Id.*, 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 in response to a user request for information in Dick 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.

ix. [7.i] without the individuals in the group having authorized production of such medical information that characterizes the group;

227. 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.

228. 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.

229. 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.

x. [7.j] and providing the produced medical information to report the characteristics of the group of individuals

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

231. Further, Dick explains that the medical information is of the type that includes characteristics of a group of individuals. Ex. 1007, [0003], [0012] ("claims handling, medication information"), [0025], [0035], [0038] ("tests ordered, results obtained").

xi. [7.k] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

232. Dick describes the HIPAA framework which normally requires express authorization from individuals to produce their medical information. Ex. 1007, [0024] ("Each request for medical information includes identification of a subject and, where required by law, a signed information release form."), [0026]. But then goes on to describe how de-identification, either under the safe harbor provision in HIPAA or the statistical de-identification section, obviates the need for such permission, so permission is not obtained in those cases. *Id.*, [0026] ("In other instances, where the law does not require a release, none is obtained."), [0044]-[0047]. Landi also describes or suggests that providing the de-identified healthcare information does not require authorization from the individuals. Ex. 1008, 1:5-18, 9:52-57, 10:17-31, 12:3-8.

233. At a minimum, it would have been obvious, in view of Dick and Landi, that authorization is not needed because the produced healthcare data is de-identified for the purpose of complying with HIPAA's alternative provisions to getting explicit authorization.

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

234. In my opinion, Dick in view of Landi and Murphy describes or suggests claim 8 for the reasons cited in 7.k above.

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

235. Dick describes or suggests that the request for healthcare information is created by identifying the members in the group for which the data is sought. Ex. 1007, Fig. 1, box 105 ("Receiving a request for medical information, the request including identification of a subject"), claim 23 ("(a) searching at least one medical information repository for medical information regarding a group of preselected individuals; (b) compiling and formatting at the site of the medical information repository the results of the search in a generic format ascertainable by requester; (c) de-identifying the results of the search; and (d) providing the compiled and de-identified results of at least one of the preselected individuals to a requester upon verification that the request is authorized.").

236. Landi also describes or suggests that the quest for healthcare information includes personal identifying information that can be encrypted to match against patient records. Ex. 1008, 4:40-47 ("the public keys . . . are used for de-identifying (encrypting) one or more items of patient identifying information (i.e., information that could be used for identifying the patient) to generate an encrypted version of the patient identifying information (referred to herein as 'Encrypted IDs)'."). Landi goes on to describe examples of such a query, including for patients in a clinical trial. *Id.*, 9:52-56 ("each time data is collected, the patient should be de-identified in the same way (e.g., it is critical to know that de-identified patient 24 from a first round of collection is also present in a second round and is still de-identified as de-identified patient 24).").

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

237. Landi describes that producing the requested healthcare information includes encrypting respective identifying data. As Landi explains, the de-identified records can be encrypted before they are produced to the central data processing system. Ex. 1008, 2:41-49, 14:26-29, 4:40-47. Moreover, it would have been obvious to a POSA that these records could also be encrypted before production from the central data processing system to a third-party in order to comply with HIPAA.

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

238. 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.

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

239. 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.

240. 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.

241. 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].

242. 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.

243. 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.

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

244. Dick describes that the user's request is for medical-related information that characterizes the group of individuals in the request. 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.").

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

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

xix. [1.b] wherein the request contains identifying data for each member in the group of multiple individuals;

246. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.b for the reasons cited in 7.a and 7.d of ground 3 above.

xx. [1.c] producing, by a hardware processor, a request token for each individual included in the group of individuals;

247. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.c for the reasons cited in 7.e of ground 3 above.

xxi. [1.d] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

248. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.d for the reasons cited in 7.f of ground 3 above.

xxii. [1.e] producing the requested medical information,

249. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.e for the reasons cited in 7.g of ground 3 above.

xxiii. [1.f] the identity of the individuals being unattainable from the produced medical information

250. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.f for the reasons cited in 7.h of ground 3 above.

xxiv. [1.g] without the individuals in the group having authorized production of Such medical information that characterizes the group; and

251. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.g for the reasons cited in 7.i of ground 3 above.

xxv. [1.h] providing the produced medical information to report the characteristics of the group of individuals

252. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.h for the reasons cited in 7.j of ground 3 above.

xxvi. [1.i] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

253. In my opinion, Dick in view of Landi and Murphy describes or suggests 1.i for the reasons cited in 7.k of ground 3 above.

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

254. In my opinion, Dick in view of Landi and Murphy describes or suggests 2 for the reasons cited in 8 of ground 3 above.

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

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

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

256. In my opinion, Dick in view of Landi and Murphy describes or suggests 4 for the reasons cited in 10 of ground 3 above.

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

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

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

258. In my opinion, Dick in view of Landi and Murphy describes or suggests 6 for the reasons cited in 13 of ground 3 above.

xxxii. [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:

259. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.a for the reasons cited in 7.a of ground 3 above.

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

260. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.b for the reasons cited in 7.b and 7.c of ground 3 above.

xxxiv. [14.c] wherein the request contains identifying data for each member in the group of multiple individuals;

261. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.c for the reasons cited in 7.d of ground 3 above.

xxxv. [14.d] producing, by a hardware processor, a request token for each individual included in the group of individuals;

262. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.d for the reasons cited in 7.e of ground 3 above.

xxxvi. [14.e] comparing the request tokens to tokens associated with the de-identified medical data to find matching tokens;

263. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.e for the reasons cited in 7.f 3 above.

xxxvii. [14.f] producing the requested medical information

264. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.f for the reasons cited in 7.g of ground 3 above.

xxxviii. [14.g] with the identity of the individuals being unattainable from the produced medical information

265. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.g for the reasons cited in 7.h of ground 3 above.

xxxix. [14.h] without the individuals in the group having authorized production of Such medical information that characterizes the group; and

266. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.h for the reasons cited in 7.i of ground 3 above.

xl. [14.i] providing the produced medical information to report the characteristics of the group of individuals

267. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.i for the reasons cited in 7.j of ground 3 above.

xli. [14.j] absent authorization from the individuals in the group of individuals identified in the request containing identifying data.

268. In my opinion, Dick in view of Landi and Murphy describes or suggests 14.j for the reasons cited in 7.k of ground 3 above.

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

269. In my opinion, Dick in view of Landi and Murphy describes or suggests 15 for the reasons cited in 8 of ground 3 above.

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

270. In my opinion, Dick in view of Landi and Murphy describes or suggests 16 for the reasons cited in 9 of ground 3 above.

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

271. In my opinion, Dick in view of Landi and Murphy describes or suggests 17 for the reasons cited in 10 of ground 3 above.

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

272. In my opinion, Dick in view of Landi and Murphy describes or suggests 18 for the reasons cited in 11 of ground 3 above.

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

273. In my opinion, Dick in view of Landi and Murphy describes or suggests 19 for the reasons cited in 13 of ground 3 above.

xlvii. [20] The computer readable storage devices of claim 14, in which producing the requested medical information includes

determining if the group includes at least a minimum number of individuals.

274. In my opinion, Dick in view of Landi and Murphy describes or suggests 20 for the reasons cited in 12 of ground 3 above.

X. CONCLUSION

275. 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 ’685 Patent to be both well-known and obvious at the time of the purported invention.

276. 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.

277. 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.

Dec. 3, 2023

Paul C. Clark DSc.