Trial Tokenization Accelerating Innovation in SEPRA – A Pragmatic Randomized Trial

https://sciencehub.novonordisk.com/ispor-us2023/Zacherle.html?cid=qr-7368761839

Zacherle E[1], Nordahl H[1], Morgan J[2], Liang M[1], Leonard S[2]

Poster SA16

Introduction

- The SEmaglutide once-weekly randomized PRAgmatic Trial (SEPRA) is an ongoing, 2-year, Phase IV trial comparing once-weekly subcutaneous semaglutide with other standard of care treatments for type 2 diabetes in routine clinical practice in the US.
- Traditionally, endpoints such as healthcare resource utilization and adherence are patient-reported in randomized trials.
- SEPRA utilized tokenization and linkage of study participants' trial data to their claims data to support evaluation of supportive secondary endpoints on comorbidities, adherence, and healthcare resource utilization.
- The objectives of this analysis were to evaluate the overlap of participants with available claims data and compare baseline characteristics with the overall trial population.

Methods

- Two closed claims datasets were used for linkage to support longitudinal analyses in the trial. The first covers >150 unique payers and >120 million patients in the US. The second dataset, by Evernorth, is a multi-payer dataset covering >35 million patients.
- Upon completion of informed consent, HealthVerity's privacy-preserving record linkage technology de-identified SEPRA participants' personally identifiable data with a unique, persistent token.
- Tokens were matched within HealthVerity's real-world data ecosystem to evaluate the number of participants with available closed medical and/or pharmacy claims data 1-year prior to randomization between July 1, 2017 and October 31, 2020 (**Figure 1**).
- Baseline characteristics were recorded in the 1-year pre-randomization period and were compared against the total trial population.

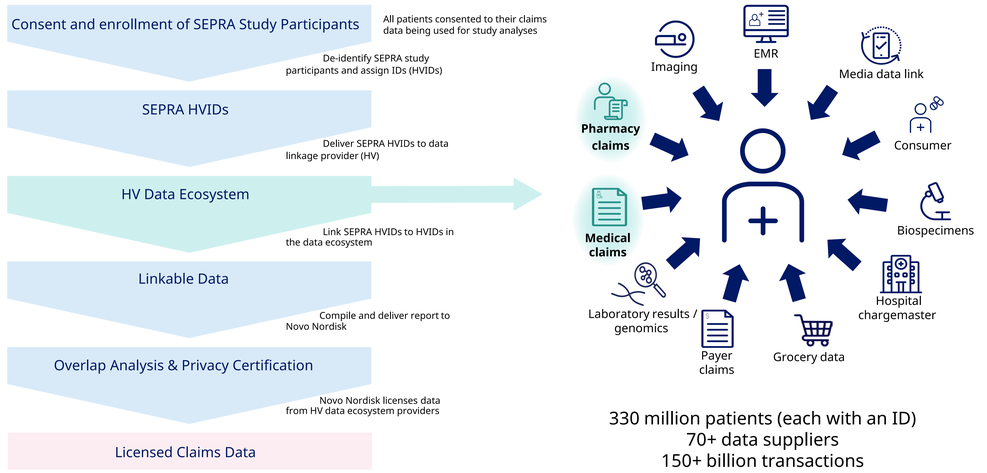
Results

- Data from 1,107 (87%) participants were tokenized and 633 (49.5%) had any claims data available during the licensed data time period.
- After requiring 1-year baseline enrolment (**Figure 2**):
  - 32% had both medical and pharmacy claims data
  - 35% had medical, and
  - 37% had pharmacy claims data.
- Similar trends in baseline characteristics were apparent between the overall trial population and HealthVerity claims data, respectively (**Table 1**):
  - Mean age of participants: 57.4 vs. 55.0 years
  - 45.8 vs. 45.2% female.
- Most participants with claims data lived in the US South (46.1%), followed by the Midwest (23.2%) (**Table 1**).

What is Trial Tokenization?

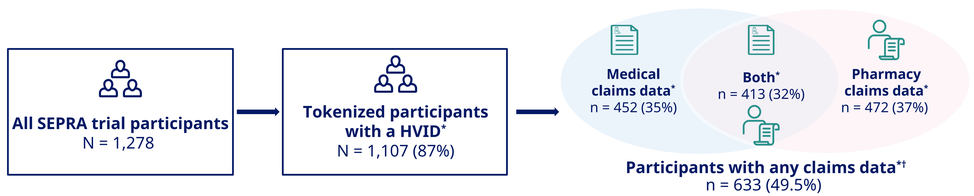
- Tokenization allows linking of disparate data sources to provide a more complete, longitudinal picture of a patient's medical history without breaching patient privacy.
- Patients' personally identifiable information are de-identified through generation of a patient-specific, encrypted 'token' (one-way de-identification) that does not compromise confidentiality.
- In clinical trials, it can enable the capture and evaluation of unique endpoints which are not regularly collected within the trial.

Figure 1: Overview of Tokenization and Data Linkage.



EMR, electronic medical records; HVID, HealthVerity identification.

Figure 2: Overlap Analysis.



Based on the licensed data period and timing of enrollment; not all patients had 1-year baseline data and were excluded from the analysis.
*Proportion of participants based on all SEPRA trial participants; †Refers to patients with pharmacy or medical claims data at any time point during the licensed data period. HVID, HealthVerity identification.

Table 1: Baseline Characteristics of Total Trial Population vs Claims Data. **Key result**

	Total trial population (N = 1,278)	HV claims data (N = 633)
Age (years), mean (SD)	57.36 (11.142)	55.04 (10.24)
Age category, n (%)		
<55 years	520 (40.7)	259 (40.9)
≥55 years	758 (59.3)	374 (59.1)
Gender, n (%)		
Male	691 (54.2)	346 (54.7)
Female	583 (45.8)	286 (45.2)
Missing	4 (0.3)	1 (0.2)
Geographic region, n (%)		
Northeast	167 (13.1)	68 (10.7)
Midwest	333 (26.1)	147 (23.2)
South	561 (43.9)	292 (46.1)
West	217 (17.0)	92 (14.5)
Unknown	0	34 (5.4)
OAD treatment, n (%)		
1 OAD	656 (51.3)	204 (43.2)*
2 OADs	459 (35.9)	157 (33.3)*

*n = 472. HV, HealthVerity; OAD, oral anti-diabetes drug; SD, standard deviation.

Conclusion

- Approximately half of the participants enrolled in SEPRA had available claims data and their baseline demographics were similar to the overall trial population.
- The use of this innovative approach within SEPRA provides an example for future research studies wherein multiple data sources are linked to evaluate patient outcomes through tokenization.
- In compliance with HIPAA, these data will be used in combination with other healthcare claims data for the planned 1-year analyses.

[1]Novo Nordisk Inc., Plainsboro, NJ, USA; [2]HealthVerity, Philadelphia, PA, USA.
This trial was sponsored by Novo Nordisk and is registered with ClinicalTrials.gov (NCT03596450).
The authors acknowledge the medical writing assistance of Laura Ward, a contract writer, and Laura Moore of Apollo, OPEN Health Communications, funded by Novo Nordisk Inc. in accordance with Good Publication Practice (GPP) guidelines. Presented at the ISPOR Annual Meeting 2023, Boston, MA, USA, May 7–10, 2023.