

1 **Title:** Does the Diabetes Alliance Program Plus Reduce Hospitalisations in Patients with Type 2  
2 Diabetes Attending Primary Care Practices? A Target Trial Protocol for Emulating a Cluster Trial  
3 Using Linked General Practice and Tertiary Health Data

4 **Authors:** Joshua Aaron Dizon<sup>1,2\*</sup>, Dr Daniel Barker<sup>1,2</sup>, Prof Shamasunder Acharya<sup>1,3</sup>, A. Prof  
5 Christopher Oldmeadow<sup>1,2</sup>, Dr Patrick Skippen<sup>1,2</sup>, Prof Alexis Hure<sup>1,2</sup>

6 <sup>1</sup>School of Medicine and Public Health, College of Health, Medicine and Wellbeing, University of  
7 Newcastle, Callaghan, NSW, 2308, Australia

8 <sup>2</sup>Hunter Medical Research Institute, New Lambton Heights, NSW, 2305, Australia

9 <sup>3</sup>Hunter New England Local Health District, John Hunter Hospital, NSW, Lookout Road, New  
10 Lambton Heights 2305, Australia

11 \* Corresponding author.

12 Joshua Aaron Dizon

13 Data Sciences

14 Hunter Medical Research Institute

15 New Lambton Heights, NSW 2305, Australia

16 Email: Joshua.Dizon@newcastle.edu.au

## 17 **ABSTRACT**

### 18 **Background:**

19 Effective management of type 2 diabetes is an increasing challenge for health services  
20 globally. Integrated care, targeted at increasing capacity in primary care for earlier  
21 intervention in type 2 diabetes, may reduce adverse health outcomes, with benefits accrued  
22 over time. This protocol aims to describe a causal approach for the evaluation of a specialist-  
23 led integrated model of care delivered in Australian general practices, using linked  
24 administrative health data.

### 25 **Methods:**

26 This protocol outlines an observational cohort study using the Lumos data asset (New South  
27 Wales Health), linking general practice and hospital data. It follows the target trial  
28 framework, emulating a parallel cluster-randomised trial, and applies the estimands  
29 framework proposed by the International Council for Harmonisation of Technical  
30 Requirements for Pharmaceuticals for Human Use. It describes the data structure and  
31 statistical analysis plan required for causal inference. The primary causal estimand is the  
32 effect of the Diabetes Alliance Program Plus, an integrated care intervention, on all-cause  
33 hospitalisation rates for adults actively attending general practice with current or newly  
34 diagnosed type 2 diabetes over a 5-year follow-up period.

### 35 **Discussion:**

36 This protocol applies the target trial framework to emulate a parallel cluster-randomized trial  
37 of the Diabetes Alliance Program Plus using linked administrative data. This approach  
38 addresses confounding and selection bias inherent in observational evaluations of practice-  
39 level interventions and will generate robust real-world evidence to inform policy regarding  
40 type 2 diabetes management in primary care settings.

41 **Trial registration:**

42 ACTRN12622001438741; 10th November 2022, retrospectively registered:

43 <https://www.anzctr.org.au/ACTRN12622001438741.aspx>.

44 **INTRODUCTION**

45 Diabetes is a major global health challenge, contributing significantly to disability, premature  
46 death, and economic burden across all regions(1). In Australia, approximately 1.3 million of  
47 people were living with type 2 diabetes mellitus (T2DM) in 2021, with higher prevalence in  
48 remote and socioeconomically disadvantaged areas(2). It is estimated that up to 500000  
49 remain undiagnosed, increasing the risk of complications and premature mortality(3).

50 T2DM is associated with a wide range of long-term health complications including  
51 cardiovascular disease, chronic kidney disease, retinopathy, neuropathy, and psychological  
52 distress(4–6). These complications reduce quality of life, increase demand on health  
53 services(2), and increase healthcare costs(7). However, early diagnosis and appropriate  
54 escalations in treatment can prevent or delay these outcomes.

55 Despite the central role of primary care in diabetes management in Australia, an estimated  
56 37% of patients do not receive appropriate care for diabetes from general practitioners(8).  
57 Barriers include limited access to decision-support tools, inadequate clinical information  
58 systems, and insufficient opportunities for continuing professional education(9). These  
59 challenges underscore the need for system-level interventions that strengthen primary care  
60 capability and capacity, through integrated care pathways, improved digital infrastructure,  
61 and stronger linkages with specialist and community services.

62 One such initiative is the Diabetes Alliance Program Plus (DAP+), an integrated care model  
63 originally launched as the Diabetes Alliance Program in 2015 as a partnership between the  
64 Hunter New England Local Health District and Hunter New England and Central Coast

65 Primary Health Network(10). The program expanded to DAP+ in 2023 to include the  
66 University of Newcastle and Hunter Medical Research Institute. DAP+ aims to build general  
67 practice capability and capacity in the Hunter New England Local Health District, New South  
68 Wales through a quality-improvement program including: (i) whole-practice data analysis  
69 and feedback; (ii) specialist-led case conferencing; and (iii) structured clinician  
70 education(10,11)..

71 Early evaluations of DAP+ demonstrated significant short-term improvements in glycaemic  
72 control (HbA1c), systolic blood pressure, and cardiovascular risk among participating  
73 patients, and the program was well received by both providers and patients(10,12,13). These  
74 findings suggest that DAP+ enhances the quality and consistency of diabetes care in general  
75 practice by supporting data-driven monitoring, multidisciplinary collaboration, and proactive  
76 case management. Improvements in these intermediate markers are clinically meaningful, yet  
77 the durability of these effects and their translation into long-term outcomes remain uncertain.  
78 Sustained improvements through continued health service engagement could reduce diabetes-  
79 related complications, hospitalisations, and premature mortality, with corresponding benefits  
80 for health system efficiency and patient wellbeing.

81 Although randomised controlled trials (RCTs) are the gold standard for assessing intervention  
82 effectiveness, they are often impractical in real-world settings. DAP+ was not designed as an  
83 RCT and has since been widely implemented as routine practice across the Hunter New  
84 England Local Health District. Therefore, this study will apply the target trial emulation  
85 framework(14), which enables causal inference from observational data.

86 This protocol outlines an emulated trial to be conducted using the Lumos data asset (NSW  
87 Health) to evaluate the long-term effectiveness of DAP+ on health service use and health  
88 outcomes. Specifically, it aims to estimate the causal effect of the DAP+ cluster intervention

89 on all-cause hospital admissions (primary outcome), emergency department presentations,  
90 hospital length of stay, lower-limb loss, adherence to testing guidelines, and biomarker levels.  
91 It will also determine the feasibility of using the Lumos data asset for evaluating pragmatic,  
92 system-level interventions delivered in primary care.

## 93 **METHODS AND ANALYSIS**

### 94 **Target trial emulation**

95 Target trial emulation applies the design principles of RCTs to observational data to  
96 strengthen causal inference and reduce bias. The protocol of a target trial specifies eligibility  
97 criteria, treatment strategies, assignment procedures, follow-up, outcomes, and the statistical  
98 analysis plan. It also defines the causal contrast of interest, typically representing the  
99 observational analogues of intention-to-treat and per-protocol effects. In this study, the target  
100 trial framework is used to emulate a cluster trial using the Lumos data asset, with all elements  
101 of trial design, including eligibility, treatment strategy, assignment, follow-up, outcomes,  
102 causal contrast, and analysis, defined *a priori* to minimise bias and enhance reproducibility.

### 103 **Data sources and structure**

104 The Lumos data asset is an ethically approved data linkage program led by the New South  
105 Wales Ministry of Health, in partnership with Primary Health Networks and the Centre for  
106 Health Record Linkage (CheReL)(15). Lumos securely links general practice data with  
107 multiple New South Wales (NSW) health data collections, including hospital admission,  
108 emergency department, ambulance, non-admitted patient and mortality records, which  
109 enables a comprehensive view of patients' health care access and outcomes. Participating  
110 practices provide informed consent for patient records to be linked with statewide health  
111 service data, which is enabled through a waiver of individual patient consent under the  
112 Lumos Governance Framework(16). All patient and practice level data are deidentified.

113 The Lumos dataset undergoes regular data quality assessments to ensure completeness,  
114 accuracy, and consistency across contributing general practices and linked health system  
115 datasets. Previous evaluations have shown the Lumos cohort to be broadly representative of  
116 the NSW population with respect to key demographic and clinical characteristics(17).

117 Use of the Lumos data asset is ethically approved to support public health programs in NSW,  
118 including chronic disease management and early intervention. Authorised collaborators  
119 access data through the cloud-based Secure Analytics Health Environment (SAPHE)  
120 platform. Ethical use of Lumos data is restricted to planning, funding, management, and  
121 evaluation of health services(16).

## 122 Target trial specification

123 A comparative summary of the ideal randomised trial and its emulated target trial is presented  
124 in [**Table 1 Placeholder**]

## 125 *Population and eligibility*

126 The target population includes people with a clinical diagnosis of T2DM who attend general  
127 practices participating in the Lumos data linkage program, within NSW. Eligibility is  
128 determined dynamically. Patients aged 18 years or older with a current T2DM diagnosis and  
129 an active patient status are eligible. Patients not meeting these criteria at a given time may  
130 become eligible once the criteria are satisfied. The eligibility definition captures both  
131 prevalent and incident T2DM diagnoses within the Lumos data asset.

132 Baseline covariates will include demographic, socioeconomic, and comorbidity measures  
133 defined at the index date, and frequency of visits to general practices, and prior  
134 hospitalisations up to two years preceding the index date. Patients must have complete  
135 baseline data on all variables required for propensity score estimation.

136 *Recruitment and follow-up period*

137 Recruitment to the target trial will begin on 1 January 2015 and continue until one year prior  
138 to the latest hospital admission record date in the Admitted Patient Data Collection. This  
139 ensures that all eligible individuals contribute at least one year of potential follow-up within  
140 the five-year outcome window.

141 Eligible patients from DAP+ intervention practices will enter the cohort on the date of their  
142 first eligible visit (index date) after participating in the program. Propensity-matched controls  
143 will be selected from eligible patients attending non-DAP+ practices outside the Hunter New  
144 England Local Health District, to minimise intervention contamination when practitioners  
145 change practices. Outcomes will be evaluated for three fixed intervals relative to the index  
146 date: baseline to 1 year, years 2–3, and years 4–5.

147 Follow-up will begin at the index date and continue until the earliest of (i) five-years post-  
148 index, (ii) death (ascertained from the Registry of Births, Deaths, and Marriages), or (iii)  
149 administrative censoring at the data cut-off date. Baseline variables for demographic,  
150 socioeconomic, and comorbidities will be measured at the index date, and health service  
151 access variables within the preceding two years of the index date.

152 *Treatment assignment procedure*

153 The treatment is DAP+ intervention, delivered in general practice as per the original clinical  
154 trial registration (ACTRN12622001438741). The comparator is standard general practice  
155 care. Exposure to treatment is defined as the first attendance at a DAP+ practice after meeting  
156 eligibility. Exposure to standard care is defined as attendance at a non-DAP+ practice after  
157 meeting eligibility.

158 Because practices are not randomly assigned, treatment assignment will be emulated using  
159 propensity score matching to balance measured confounders between exposure groups. The

160 resulting matched cohort will approximate randomisation conditional on observed covariates.  
161 Details of the propensity score estimation, matching specification, and balance diagnostics  
162 are described in the *Primary Analyses* section.

### 163 *Outcomes*

164 The primary outcome is total all-cause hospitalisations, defined as the number of unique  
165 patient episodes of care. All-cause hospitalisations will be identified in from the New South  
166 Wales Admitted Patient Data Collection (APDC). Multiple same-day admissions will be  
167 counted as one episode of care to obtain unique per patient per day events. Totals will be  
168 separately calculated for each follow-up interval (baseline to 1 year, years 2-3, and years 4-  
169 5).

170 Secondary outcomes include:

- 171 i. hospital length of stay: number of days from admission start date to discharge  
172 date as recorded in the APDC. When multiple admissions occur on the same date,  
173 the episode with the longest length of stay is used.
- 174 ii. potentially preventable hospitalisations: defined according to Australian Institute  
175 of Health and Welfare indicators for selected potentially preventable  
176 hospitalisation (Table S4 in Additional file 1), with the same definition of episode  
177 of care as applied to the primary outcome.
- 178 iii. T2DM-specific hospitalisations, defined as a hospitalisation with T2DM  
179 diagnosis recorded as the principal diagnosis in the APDC. ICD-10 clinical codes  
180 for determining T2DM diagnoses is available in Table S1 in Additional file 1.
- 181 iv. ED presentations: identified using the Emergency Department Data Collection  
182 (EDDC), with the same definition of episode of care as applied to the primary  
183 outcome.

184

185 v. potentially avoidable ED presentations: defined according to the Australian  
186 Institute of Health and Welfare indicators for selected potentially avoidable GP-  
187 type presentations to emergency departments (Table S5 in Additional file 1), with  
188 the same definition of episode of care as applied to the primary outcome.

189 vi. T2DM-specific ED presentations, defined as ED presentations with T2DM  
190 diagnosis recorded as the principal diagnosis in the EDDC. The EDDC records  
191 diagnosis codes using the ICD-9, ICD-10, or SNOMED-CT code sets. Codes  
192 used to identify T2DM-specific ED presentations in each system are listed in  
193 Tables S2 and S3 in Additional file 1.

194 vii. T2DM-related lower-limb loss: defined according to the Australian Institute of  
195 Health and Welfare indicator for Hospitalisation for lower limb amputation with  
196 T2DM as principal or additional diagnosis (Table S6 in Additional file 1).

197 viii. all-cause mortality: death dates will be obtained from Registry of Births, Deaths,  
198 and Marriages data set.

199 ix. Adherence to testing frequency guidelines for T2DM pathology monitoring:  
200 testing frequency will use definitions from the Royal Australian College of  
201 General Practitioners' annual cycle of care guidelines for T2DM(18). Separate  
202 adherence measures will be constructed for HbA1c, lipid profile, and renal  
203 function testing. For HbA1c, each test initiates a 6-month coverage window  
204 extending from test date. For lipid profile and renal function tests, each test  
205 initiates a 12-month coverage window. If a subsequent test occurs within an  
206 active coverage window, the coverage is extended from the new test date for the  
207 relevant duration. Coverage from tests conducted before the start of a follow-up  
208 interval is carried forward into the interval. Coverage intervals that extend

209 beyond the end of a follow-up interval will be truncated at the interval's end. For  
210 each test type, the proportion of months within the interval covered by any active  
211 window is calculated per participant. Participants with more than 90 percent  
212 coverage for a given test type are classified as adherent to that guideline.

213 x. T2DM-related biomarker levels (HbA1c, lipid profile, eGFR): Laboratory  
214 measurements will be obtained from General Practice Electronic Health Records.  
215 For each biomarker, the mean value will be calculated separately for each patient  
216 within each follow-up interval. When multiple tests occur within an interval, all  
217 available results will contribute to the per patient interval mean. Analyses will be  
218 performed separately for each biomarker. Additionally, biomarkers levels may be  
219 classified according to clinically relevant thresholds (e.g., HbA1C <7%, 7-8.9%,  
220  $\geq 9\%$  etc.).

## 221 Clarifying the estimands

222 An estimand defines the treatment effect that a study aims to quantify. Clear specification of  
223 estimands clarifies the research question and guides appropriate methodological choices.

224 Estimands in this protocol are defined according to the International Council for  
225 Harmonisation (ICH) E9(R1) guidelines and the framework proposed by Kahan (2024)(19) .

226 The matched population will be used to estimate the intention-to-treat (ITT) effect. Handling  
227 of intercurrent events, which are post-baseline events that may affect interpretation, will also  
228 be specified. The corresponding estimands are summarised in Table 2.

229 **[Table 2 Placeholder]**

## 230 Primary Analyses

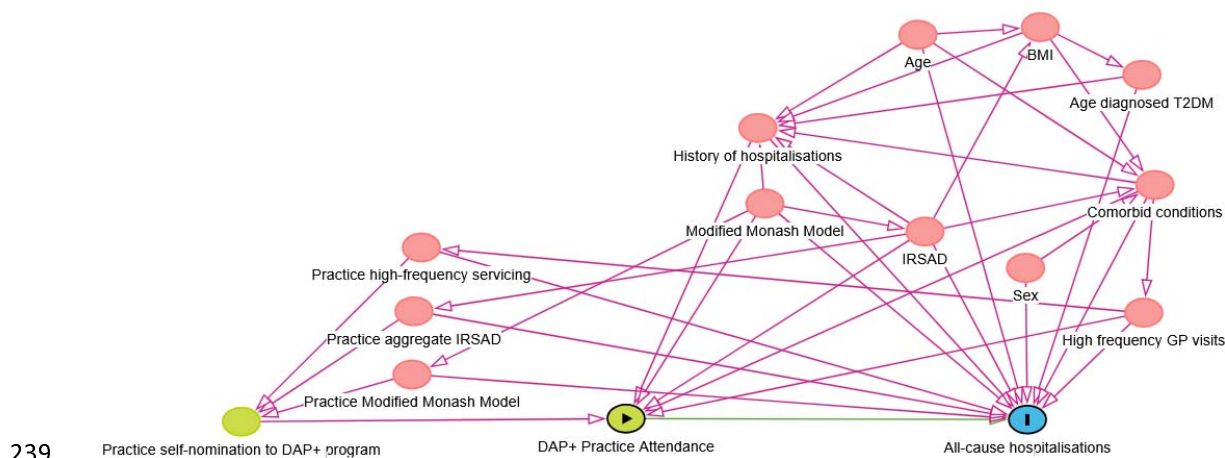
### 231 General principles

232 All analyses will comply with Lumos data governance requirements, including suppression of  
233 cells with counts  $\leq 5$ , which will be reported as ' $\leq 5$ '.

### 234 Causal analysis principles

235 An observational analogue of the ITT effect will be applied to the primary analysis. DAP+  
236 exposure will be compared to the relevant comparator (standard care) using a propensity  
237 matched cohort.

238 *Figure 1. Directed Acyclic Graph representing relations between DAP+ exposure, outcome, and confounders*



240 Relevant covariates for creating the propensity matched cohort were identified using the  
241 directed acyclic graph (DAG) in Figure 1, which was informed by literature review(20–23)  
242 and clinical domain knowledge. It was determined that treatment assignment is  
243 predominantly affected by geographic factors and general practice characteristics.

244 Access to the DAP+ enhanced care intervention depends not only on individual  
245 characteristics but also on whether a patient's general practice self-nominated to participate in  
246 DAP+. Individual exposure can occur only within participating practices. Self-nomination  
247 may be influenced by the implementation of DAP+, which primarily targets, though is not

248 limited to, regional general practices(10). Practice-level aggregates of Index of Relative  
249 Socioeconomic Advantage and Disadvantage (IRSAD, an Australian national measure  
250 ranking areas by relative socioeconomic conditions based on income, education, and  
251 employment)(24), geographic classification under the Australian Modified Monash Model  
252 (MMM, which categorises areas by remoteness and population size)(25), and operational  
253 features such as high-frequency servicing (>30% of patients attending  $\geq 12$  times over two  
254 years)(20) were identified as key determinants of participation. Because these factors are also  
255 associated with health outcomes, they form cluster-level biasing pathways upstream of the  
256 exposure–outcome relationship.

257 At the individual level, factors influencing general practice attendance include  
258 sociodemographic status via IRSAD, geographic remoteness via MMM, comorbidities,  
259 frequency of general practice visits, and prior hospitalisations. In Australia, multimorbidity is  
260 more prevalent in regional and remote areas than urban settings(26). As these factors are  
261 associated with both health outcomes and the likelihood of attending practices in different  
262 socioeconomic or rural settings, they may also form biasing pathways between the  
263 intervention and outcomes. Although BMI, age, and sex lie on the causal pathway, they are  
264 not required in the minimal adjustment set. Adjustment for individual MMM, IRSAD, and  
265 comorbid conditions is sufficient to close the identified individual-level biasing paths in the  
266 DAG.

#### 267 Descriptive statistics

268 Summary statistics of baseline characteristics will be stratified by treatment group and will  
269 include where available in the Lumos data asset; the age at study entry, sex, BMI, IRSAD,  
270 MMM, comorbidities, and health service access indicators. Summary measures will be mean

271 and standard deviation (SD) for continuous variables, and frequencies and percentages for  
272 categorical variables.

### 273 Balancing cluster and individual confounders via propensity scores

274 The minimal adjustment set required to identify the causal effect of receiving the intervention  
275 includes practice-level MMM, aggregate IRSAD, high-frequency servicing indicator, and  
276 individual-level MMM, IRSAD, comorbidities, and health service use indicators. To improve  
277 the precision of estimates, individual BMI, sex, and age will also be included as additional  
278 covariates.

279 Before individual-level matching, intervention and control practices will be exactly matched  
280 on MMM classification, aggregate IRSAD, and high-frequency servicing status. Each  
281 intervention cluster may be paired with one or more control clusters to balance cluster-level  
282 confounding on these variables.

283 After balancing cluster-level confounders by forming cluster-similar groups, intervention-  
284 exposed individuals will be matched to control individuals within these groups. Individual  
285 matching will be based on propensity scores estimated using a logistic regression model, with  
286 individual exposure to the intervention at baseline as the outcome and individual level  
287 confounders as predictors. Each DAP+ participant will be matched to the control with the  
288 closest propensity score who did not attend a DAP+ practice at baseline. Matches will have  
289 propensity scores within 0.2 pooled standard deviations of the logit of the propensity score to  
290 improve covariate balance between matched pairs(27). Matching will be performed one-to-  
291 one without replacement, and unmatched individuals will be excluded.

292 Covariate balance will be assessed in terms of standardised mean differences after propensity  
293 matching to determine whether balancing was successful. Interactions of variables and non-

294 linear functional forms of continuous variables will be considered to iteratively improve  
295 covariate balance, as necessary.

### 296 *Outcome analysis*

297 All outcomes will be analysed using longitudinal Bayesian hierarchical models to account for  
298 clustering and repeated measurements within individuals and dyad pairs. Models will be fit  
299 using the propensity-matched sample. All models will include fixed effects for treatment,  
300 time, and their interaction, and random intercepts to account for repeated measures within  
301 clusters and dyad pairs. Weakly informative priors will be specified for all fixed and random  
302 effects to improve model regularisation and ensure stable Markov chain Monte Carlo  
303 sampling without imposing restrictive assumptions.

### 304 *Primary outcome*

305 The primary outcome is the total number of hospitalisations in all time intervals. For each  
306 patient, all hospitalisations within an interval will be summed to obtain a single patient-level  
307 total. This outcome will be modelled using a Bayesian hierarchical Poisson, negative-  
308 binomial, or zero-inflated Poisson model, as selected by model fit, evidence of  
309 overdispersion, or excess zeroes. The model will estimate incidence rate ratios (IRR) with  
310 95% credible intervals (CrI).

### 311 *Secondary outcomes*

312 Secondary outcomes include total hospital length of stay, total potentially preventable  
313 hospitalisations, total T2DM-specific hospitalisations, total ED presentations, total potentially  
314 avoidable ED presentations, and T2DM-specific ED presentations in all time intervals. Each  
315 will be analysed using the same modelling considerations as the primary outcome to estimate  
316 IRR with 95% CrI. Statewide data indicate low event rates for T2DM-specific  
317 hospitalisations (3.7%)(2), raising the possibility of nonconvergent models in the matched

318 cohort. If models fail to converge, results will be reported descriptively by treatment group  
319 and time interval. Similar considerations apply to T2DM-specific ED presentations.

320 T2DM-related lower limb loss will be analysed using a logistic regression model. Posterior  
321 predictions of event probabilities will be used to estimate relative risks (RR) with 95% CrI.

322 All-cause mortality will be analysed using a pooled logistic regression model for discretised  
323 time-to-event. Cumulative survival probabilities will be obtained via posterior predictions to  
324 estimate risk differences (RD) with 95% CrI at 1, 3, and 5 years.

325 Pathology biomarker outcomes will be analysed using normal or log-normal regression  
326 models based on empirical distribution of each biomarker. For biomarkers with clinically  
327 informative thresholds, secondary analyses will be conducted using logistic, multinomial, or  
328 ordinal regression models where appropriate.

329 Compliance with annual cycle-of-care testing for metabolic markers will be analysed using  
330 logistic regression models.

## 331 **DISCUSSION**

332 Evidence indicates that the risk of hospital admission among patients with type 2 diabetes  
333 mellitus (T2DM) is associated with elevated HbA1c levels(28,29) and other routinely  
334 measured primary care markers, including increased uACR, elevated systolic blood pressure,  
335 and reduced high-density lipoprotein (HDL) cholesterol(29). The association between  
336 elevated HbA1c and macrovascular and microvascular complications has been well  
337 established since 2000, with each 1% HbA1c reduction associated with 21% lower diabetes-  
338 related endpoints and 37% lower microvascular complication (30),underscoring the  
339 importance of metabolic control in preventing hospitalisations and other adverse outcomes. In  
340 the pilot evaluation of DAP+, improvements in diabetes-related metabolic markers,

341 particularly HbA1c, were observed among T2DM patients attending DAP+-enabled general  
342 practices(10,11). Given the established link between these markers and diabetes-related  
343 complications, these findings suggest that DAP+ may contribute to improved long-term  
344 outcomes for patients with T2DM through enhanced metabolic control. However, as the pilot  
345 study was not designed to assess long-term endpoints, the effect of DAP+ on longer-term  
346 outcomes remains an important question warranting further investigation. Potentially  
347 preventable hospitalisations are widely used to monitor the performance of primary health  
348 care systems and to evaluate the effectiveness of primary care interventions in Australia and  
349 internationally(31). As the DAP+ program aims to strengthen chronic disease management in  
350 general practice, reductions in potentially preventable hospitalisation rates represent a  
351 plausible downstream indicator of its effectiveness in improving diabetes care and long-term  
352 health outcomes.

353 Ideally, the effectiveness of an intervention would be assessed using a RCT. However, as the  
354 DAP+ program has been progressively implemented across the Hunter and New England  
355 region since 2015, conducting a new RCT in this setting is not feasible due to potential  
356 contamination and spillover from prior program exposure, and the cost of running such a  
357 large trial. Implementing an RCT in other health districts would also pose substantial  
358 practical and resource challenges, as it would require establishing the DAP+ model of care de  
359 novo within a different health service context. There are also ethical considerations, given the  
360 existing evidence supporting its intermediate benefits (REFS), making it questionable to  
361 withhold access to enhanced care. Furthermore, in regions without the DAP+ model, patients  
362 have limited access to metropolitan specialist care and remain at increased risk of premature  
363 and progressive diabetes-related complications due to undertreated disease.

364 DAP+ has been progressively implemented across the Hunter New England Local Health  
365 District in partnership with the Hunter New England and Central Coast Primary Health  
366 Network since 2015, and administrative health data from participating general practices are  
367 routinely captured through the Lumos data linkage program in New South Wales(17). This  
368 linkage provides a rich longitudinal dataset suitable for evaluating the long-term effects of  
369 DAP+. Using these existing data within a target trial emulation framework(14) offers a  
370 practical and robust alternative to conducting a new RCT. Previous studies have shown that  
371 well-designed emulated trials using observational data can yield treatment effect estimates  
372 consistent with those obtained from true RCTs(32).

373 Although this protocol employs propensity score methods to balance key confounders,  
374 residual confounding cannot be fully eliminated and remains a limitation of any observational  
375 study(33–36). Unmeasured differences may exist across Local Health Districts and Primary  
376 Health Networks in which practices are nested, potentially influencing diabetes management.  
377 However, all operate under New South Wales Health governance with uniform policy,  
378 funding, and clinical frameworks. Matching within geographic strata (metropolitan, regional,  
379 rural, remote) further limits systematic region-level bias.

380 Lumos participation is voluntary among general practices, introducing potential selection bias  
381 in the data collections. As of November 2025, more than 7 million people across NSW are  
382 represented with approximately one-third of all general practices having consented to data-  
383 sharing across the state (Add Ref). The 2023 Lumos evaluation reported 91% demographic  
384 representativeness to the broader New South Wales population(37), providing reasonable  
385 assurance that the Lumos data set captures statewide demographic diversity.

386 While recognising the limitations inherent to observational analyses, this study will generate  
387 the best available evidence to inform general practice-centred models of care for type 2  
388 diabetes mellitus in rural and remote communities in NSW.

## 389 **DECLARATIONS**

390 **Ethics approval and consent to participate:** The Diabetes Alliance Program Plus has ethics  
391 approval from the Hunter New England Human Research Ethics Committee  
392 (2024/ETH01649), registered with the University of Newcastle (R-2024-0073). The Lumos  
393 data asset has ethics approval from the Population Health Services Research Ethics  
394 Committee (2019/ETH00660) with a waiver of individual participant consent for approved  
395 secondary data uses.

396 **Consent for publication:** Not applicable.

397 **Availability of data and materials:** No new datasets were generated in this study.

398 **Competing interests:** None declared.

399 **Funding:** The Diabetes Alliance Program Plus (DAP+) and this work is supported by a  
400 philanthropic gift from the Colonial Foundation, administered by the Hunter Medical  
401 Research Institute. Joshua Dizon is supported by an Australian Government Research  
402 Training Program (RTP) scholarship administered through the University of Newcastle's  
403 Strategic Engagement Scheme.

404 **Authors' contributions:** JAD, DB, CO and AH contributed to the statistical analysis plan  
405 and statistical rationale described in this article. AH and SA provided clinical rationale for  
406 this protocol. All authors were involved in the discussions and editing of the manuscript and  
407 provide approval for publication.

408 **Patient and public involvement:** Patients and/or the public were not involved in the design,  
409 conduct, reporting, or dissemination plans of this research.

410 **Acknowledgements:** The authors acknowledge the Lumos program, a partnership of the  
411 NSW Ministry of Health and NSW Primary Health Networks made possible through the  
412 participation of general practices across NSW. Record linkage was carried out by the Centre  
413 for Health Record Linkage.

## 414 REFERENCES

- 415 1. Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, et al. Global,  
416 regional, and national burden of diabetes from 1990 to 2021, with projections of  
417 prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021.  
418 *The Lancet*. 2023 Jul 15;402(10397):203–34.
- 419 2. Australian Institute of Health and Welfare [Internet]. [cited 2025 Nov 11]. Diabetes:  
420 Australian Facts. Available from:  
421 <https://www.aihw.gov.au/reports/diabetes/diabetes/contents/summary>
- 422 3. Diabetes Australia [Internet]. [cited 2025 Nov 11]. Diabetes in Australia. Available from:  
423 <https://www.diabetesaustralia.com.au/about-diabetes/diabetes-in-australia/>
- 424 4. Galiero R, Caturano A, Vetrano E, Beccia D, Brin C, Alfano M, et al. Peripheral  
425 Neuropathy in Diabetes Mellitus: Pathogenetic Mechanisms and Diagnostic Options. *Int*  
426 *J Mol Sci*. 2023 Feb 10;24(4).
- 427 5. Dal Canto E, Ceriello A, Rydén L, Ferrini M, Hansen TB, Schnell O, et al. Diabetes as a  
428 cardiovascular risk factor: An overview of global trends of macro and micro vascular  
429 complications. *Eur J Prev Cardiol*. 2019 Dec;26(2\_suppl):25–32.
- 430 6. Akshatha S, Nayak UB. The psychological distress associated with type 2 diabetes  
431 mellitus represents an unmet need for drug discovery. *Medicine in Drug Discovery*.  
432 2024;23:100196.
- 433 7. Lee CMY, Colagiuri R, Magliano DJ, Cameron AJ, Shaw J, Zimmet P, et al. The cost of  
434 diabetes in adults in Australia. *Diabetes Research and Clinical Practice*. 2013  
435 Mar;99(3):385–90.
- 436 8. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al.  
437 CareTrack: assessing the appropriateness of health care delivery in Australia. *Medical*  
438 *Journal of Australia*. 2012 Jul;197(2):100–5.
- 439 9. Brown JB, Harris SB, Webster-Bogaert S, Wetmore S, Faulds C, Stewart M. The role of  
440 patient, physician and systemic factors in the management of type 2 diabetes mellitus.  
441 *Family Practice*. 2002 Aug 1;19(4):344–9.

- 442 10. Acharya S, Philcox AN, Parsons M, Suthers B, Luu J, Lynch M, et al. Hunter and New  
443 England Diabetes Alliance: innovative and integrated diabetes care delivery in general  
444 practice. *Australian Journal of Primary Health*. 2019;25(3):219.
- 445 11. Acharya S, Taylor R, Parsons M, Attia J, Leigh L, Oldmeadow C, et al. Spillover effects  
446 from a type 2 diabetes integrated model of care in 22,706 Australians: an open cohort  
447 stepped wedge trial. *BMC Endocr Disord*. 2024 Sep 10;24(1):183.
- 448 12. Harris ML, Kuzulugil D, Parsons M, Byles J, Acharya S. “They were all together ...  
449 discussing the best options for me”: Integrating specialist diabetes care with primary care  
450 in Australia. *Health & Social Care in the Community*. 2021 Sep;29(5):e135–43.
- 451 13. Taylor RM, Acharya SH, Parsons ME, Ranasinghe UP, Kuzulugil DO, Fleming KC, et  
452 al. Australian practice nurses’ perspectives on integrating specialist diabetes care with  
453 primary care: a qualitative study. *Fam Pract*. 2025 Apr 12;42(3):cmf020.
- 454 14. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized  
455 Trial Is Not Available: Table 1. *American Journal of Epidemiology*. 2016 Apr  
456 15;183(8):758–64.
- 457 15. NSW Health [Internet]. [cited 2025 Nov 11]. About Lumos. Available from:  
458 <https://www.health.nsw.gov.au/lumos/Pages/about.aspx>
- 459 16. NSW Health [Internet]. 2025 [cited 2025 Nov 11]. Lumos Data Governance Framework.  
460 Available from: [https://www.health.nsw.gov.au/lumos/Publications/data-governance-](https://www.health.nsw.gov.au/lumos/Publications/data-governance-framework.pdf)  
461 [framework.pdf](https://www.health.nsw.gov.au/lumos/Publications/data-governance-framework.pdf)
- 462 17. Correll P, Feyer AM, Phan PT, Drake B, Jammal W, Irvine K, et al. Lumos: a statewide  
463 linkage programme in Australia integrating general practice data to guide system  
464 redesign. *Integrated Healthcare Journal*. 2021 May 17;3(1):e000074.
- 465 18. The Royal Australia College of General Practitioners [Internet]. 2020. Management of  
466 type 2 diabetes: A handbook for general practice.
- 467 19. Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer  
468 on the ICH E9(R1) addendum. *BMJ*. 2024 Jan 23;384:e076316.
- 469 20. NSW Health [Internet]. 2021 [cited 2025 Nov 11]. General practice activity can affect  
470 hospital visits. Available from: [https://www.health.nsw.gov.au/lumos/Factsheets/gp-](https://www.health.nsw.gov.au/lumos/Factsheets/gp-activity-insight.pdf)  
471 [activity-insight.pdf](https://www.health.nsw.gov.au/lumos/Factsheets/gp-activity-insight.pdf)
- 472 21. Phillips A. Health status differentials across rural and remote Australia. *Australian*  
473 *Journal of Rural Health*. 2009 Feb;17(1):2–9.
- 474 22. Britt HC, Harrison CM, Miller GC, Knox SA. Prevalence and patterns of multimorbidity  
475 in Australia. *Medical Journal of Australia*. 2008 Jul;189(2):72–7.
- 476 23. Glenister KM, Guymer J, Bourke L, Simmons D. Characteristics of patients who access  
477 zero, one or multiple general practices and reasons for their choices: a study in regional  
478 Australia. *BMC Family Practice*. 2021 Dec;22(1):2.

- 479 24. Australian Bureau of Statistics [Internet]. [cited 2025 Nov 11]. Socio-Economic Indexes  
480 for Areas (SEIFA), Australia. Available from:  
481 [https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-](https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#index-of-relative-socio-economic-advantage-and-disadvantage-irsad-)  
482 [index-of-relative-socio-economic-advantage-](https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#index-of-relative-socio-economic-advantage-and-disadvantage-irsad-)  
483 [and-disadvantage-irsad-](https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release#index-of-relative-socio-economic-advantage-and-disadvantage-irsad-)
- 484 25. Versace VL, Skinner TC, Bourke L, Harvey P, Barnett T. National analysis of the  
485 Modified Monash Model, population distribution and a socio-economic index to inform  
486 rural health workforce planning. *Australian Journal of Rural Health*. 2021 Oct;29(5):801–  
487 10.
- 488 26. Australian Institute of Health and Welfare [Internet]. [cited 2025 Nov 11].  
489 Multimorbidity in Australia. Available from: [https://www.aihw.gov.au/reports/chronic-](https://www.aihw.gov.au/reports/chronic-disease/multimorbidity-in-australia/contents/how-common-is-multimorbidity)  
490 [disease/multimorbidity-in-australia/contents/how-common-is-multimorbidity](https://www.aihw.gov.au/reports/chronic-disease/multimorbidity-in-australia/contents/how-common-is-multimorbidity)
- 491 27. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of  
492 Confounding in Observational Studies. *Multivariate Behav Res*. 2011 May;46(3):399–  
493 424.
- 494 28. Bo S, Ciccone G, Grassi G, Gancia R, Rosato R, Merletti F, et al. Patients with type 2  
495 diabetes had higher rates of hospitalization than the general population. *Journal of*  
496 *Clinical Epidemiology*. 2004 Nov;57(11):1196–201.
- 497 29. Tomlin AM, Dovey SM, Tilyard MW. Risk factors for hospitalization due to diabetes  
498 complications. *Diabetes Research and Clinical Practice*. 2008 May;80(2):244–52.
- 499 30. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association  
500 of glycaemia with macrovascular and microvascular complications of type 2 diabetes  
501 (UKPDS 35): prospective observational study. *BMJ*. 2000 Aug 12;321(7258):405–12.
- 502 31. Australian Commission on Safety and Quality in Health Care [Internet]. 2017 [cited 2025  
503 Nov 11]. A guide to the potentially preventable hospitalisations indicator in Australia.  
504 Available from: [https://www.safetyandquality.gov.au/publications-and-](https://www.safetyandquality.gov.au/publications-and-resources/resource-library/guide-potentially-preventable-hospitalisations-indicator-australia)  
505 [resources/resource-library/guide-potentially-preventable-hospitalisations-indicator-](https://www.safetyandquality.gov.au/publications-and-resources/resource-library/guide-potentially-preventable-hospitalisations-indicator-australia)  
506 [australia](https://www.safetyandquality.gov.au/publications-and-resources/resource-library/guide-potentially-preventable-hospitalisations-indicator-australia)
- 507 32. Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative, Franklin JM, Desai RJ,  
508 Feldman W, et al. Emulation of Randomized Clinical Trials With Nonrandomized  
509 Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023 Apr 25;329(16):1376.
- 510 33. Fewell Z, Davey Smith G, Sterne JAC. The Impact of Residual and Unmeasured  
511 Confounding in Epidemiologic Studies: A Simulation Study. *American Journal of*  
512 *Epidemiology*. 2007 Jun 24;166(6):646–55.
- 513 34. Howards PP. An overview of confounding. Part 1: the concept and how to address it.  
514 *Acta Obstetrica et Gynecologica Scandinavica*. 2018 Apr;97(4):394–9.
- 515 35. Verbeek JH, Whaley P, Morgan RL, Taylor KW, Rooney AA, Schwingshackl L, et al.  
516 An approach to quantifying the potential importance of residual confounding in

- 517 systematic reviews of observational studies: A GRADE concept paper. *Environment*  
518 *International*. 2021 Dec;157:106868.
- 519 36. Arah OA. Bias Analysis for Uncontrolled Confounding in the Health Sciences. *Annual*  
520 *Review of Public Health*. 2017 Mar 20;38(1):23–38.
- 521 37. NSW Health [Internet]. 2023 [cited 2025 Nov 11]. Lumos Evaluation Report 3. Available  
522 from: <https://www.health.nsw.gov.au/lumos/Publications/lumos-evaluation-report-3.pdf>
- 523

524 *Table 1: Summary of the design features of the target trial protocol for the Diabetes Alliance*  
 525 *Program Plus (DAP+) using the Lumos data asset.*

<b>Protocol component</b>	<b>Target trial specification</b>	<b>Target trial emulation</b>
<b>Eligibility criteria - Practices</b>	General practices that consent to participate in the trial and are eligible for randomization to receive the DAP+ intervention or continue usual care.	General practices that consent to data sharing with Lumos.  Intervention practices: General practices that consent to receive and have received the DAP+.  Control practices: General practices that do not receive the DAP+.
<b>Eligibility criteria - Patients</b>	Adults ( $\geq 18$ years) with a diagnosis of T2DM. Actively attend general practices, defined as $\geq 3$ visits to general practices in the past 2 years.  Exclusion criteria: Pregnant women Diabetes diagnoses other than T2DM.	Eligible patients will be identified from eligible practices in the Lumos data asset.  Pregnancy status which cannot be directly identified in Lumos data. Consequently, exclusion by pregnancy status cannot be performed.
<b>Treatment strategies</b>	1. Receive DAP+ integrated care intervention  2. Usual care.  Participants continue with their assigned treatment exposure during follow-up.	Same
<b>Assignment procedures</b>	Practices will be randomly assigned to receive DAP+ (intervention) or continue with usual care. Practices will not be blinded to the intervention.	Randomisation will be emulated via propensity score methods, conditional on covariates identified as important confounders via causal directed acyclic graphs.  Matching procedure: <ul style="list-style-type: none"> <li>• Cluster level: many-to-one matching of control to intervention practices, accounting for calendar time, aggregated patient rurality, aggregated patient socioeconomic factors, and high-servicing status</li> <li>• One-to-one matching of patients without replacement within matched clusters, accounting for calendar time, age, sex, age at first recorded T2DM diagnosis, rurality, socioeconomic status, comorbidities, frequency of GP</li> </ul>

		visits, and prior hospitalisations.
<b>Follow-up period</b>	Follow-up begins at the date following the first case conferencing in each site and continues until death, loss-to-follow-up, or administrative end of follow-up (5 years from baseline), whichever occurs first.	Rolling enrolment with calendar-time matching: treated patients matched to controls with index dates within $\pm 28$ days. Treated index date is the first general practice visit after practice intervention was initiation; control index date is an eligible general practice visit at a non-intervention practice within the matching window. Follow-up begins at index date and continues until death or administrative end of follow-up (5 years from index), whichever occurs first.
<b>Outcomes</b>	Outcomes are assessed at the individual level; at 1, 3, and 5 years from first exposure to DAP+: <ul style="list-style-type: none"> <li>• Total all-cause hospitalisations</li> <li>• Total all-cause hospital length of stay</li> <li>• Total potentially preventable hospitalisations</li> <li>• Total T2DM-specific hospitalisation</li> <li>• Total ED presentations</li> <li>• Total potentially avoidable ED presentations</li> <li>• Total T2DM-specific ED presentations</li> <li>• Proportion with T2DM-related lower limb loss</li> <li>• Mortality</li> <li>• HbA1c %, lipids markers (LDL, HDL, triglycerides, and total cholesterol), uACR, and eGFR.</li> <li>• Adherence to testing frequency guidelines for HbA1c, lipids, and kidney function tests.</li> </ul>	Same, noting that Urine Albumin-to-Creatinine Ratio (uACR) is not consistently available in Lumos.
<b>Causal contrasts</b>	Intention-to-treat: effect of assignment to DAP+ versus usual care;	Same
<b>Estimands</b>	ITT: Sample average treatment effect on the treated among eligible individuals visiting DAP+ practices (effect of at least one exposure to DAP+ enhanced care)	Same
<b>Analysis plan</b>	Summary of patient baseline characteristics and covariate balance diagnostics among matched cohort.  Hierarchical Bayesian models: <ul style="list-style-type: none"> <li>• Poisson, negative binomial, or</li> </ul>	Same

	<p>zero-inflated Poisson regression models for total all-cause hospitalisations, total all-cause hospital length of stay, total potentially preventable hospitalisations, total T2DM-specific hospitalisations, total ED presentations, total potentially avoidable ED presentations, and total T2DM-specific ED presentations.</p> <ul style="list-style-type: none"><li>• Logistic regression model for T2DM-related lower limb amputations</li><li>• Pooled logistic regression model for survival</li><li>• Logistic regression model for adherence to testing frequency guidelines for HbA1c, lipids, and kidney function tests.</li><li>• Normal/log-normal regression models for HbA1c, lipids, and uACR, and/or appropriate categorical regression models (logistic, multinomial, or ordinal) for relevant clinical thresholds</li></ul>	
--	--	--

526

527 *Table 2: Description of the estimands for testing the effectiveness of the Diabetes Alliance Program*  
 528 *Plus (DAP+) on hospitalisations compared to standard care in general practice.*

Attribute	Specification
Estimand	Sample average treatment effect on the treated for the intention-to-treat (ITT) population: effect of at least one exposure to DAP+ intervention (i.e. enhanced care)
Population	Adults ( $\geq 18$ years) with existing or newly diagnosed T2DM recorded in Lumos, who are active patients of general practices in New South Wales, Australia.
Treatment conditions	Intervention: Attendance at a DAP+ enhanced care practice. Comparator: Attendance at a non-DAP+ practice.
Endpoint	<p>Primary:</p> <ul style="list-style-type: none"> <li>• Total number of hospitalisations, counted as unique admissions per day, summed over three prespecified time intervals (baseline-year 1, years 2-3, and years 4-5). The primary estimand targets the treatment effect on the total number of hospitalisations per participant within these intervals, evaluated using a longitudinal model with time and treatment-by-time interaction terms.</li> </ul> <p>Secondary:</p> <ul style="list-style-type: none"> <li>• Total hospital length of stay*, total potentially preventable hospitalisations, total T2DM-specific hospitalisations, total emergency department presentations, total potentially avoidable emergency department presentations, total T2DM-specific ED presentations, T2DM-related lower limb loss, adherence to T2DM-related pathology monitoring guidelines, and T2DM-related biomarker levels, each assessed over the three prespecified time intervals (baseline-year 1, years 2-3, years 4-5). The estimand for each is the treatment effect on the mean value (total count, proportion, or continuous measure as appropriate) within intervals.</li> <li>• All-cause mortality, analysed as a discretised time-to-event outcome with survival status evaluated at 1, 3, and 5 years. The estimand is the treatment effect on the probability of survival at each prespecified time point.</li> </ul>
Summary Measure	<p>Count outcomes:</p> <ul style="list-style-type: none"> <li>• Incidence rate ratio comparing treatment vs control within time intervals.</li> </ul> <p>Binary outcomes:</p> <ul style="list-style-type: none"> <li>• Risk ratio comparing treatment vs control within time intervals.</li> </ul> <p>Continuous outcomes</p> <ul style="list-style-type: none"> <li>• Mean difference between treatment and control within time intervals.</li> </ul> <p>Discretised time-to-event outcomes:</p> <ul style="list-style-type: none"> <li>• Risk difference of survival at 1, 3, and 5 years.</li> </ul>

Handling of intercurrent events	<p>Loss of active attendance: Treatment policy strategy. Analyses will use exposure assigned at baseline, regardless of whether patients become inactive during follow-up.</p> <p>Treatment switching: Treatment policy strategy. Analyses will use exposure assigned at baseline, regardless of post-baseline switching between DAP+ and non-DAP+ practices.</p> <p>Death during follow-up: While-alive strategy. For all outcomes except all-cause mortality, data will be included from index date until the date of death.</p>
---------------------------------------	--

529

*\*When multiple admissions occur on the same date, the episode with the longest length of stay is used*

