Overview of the Analytic Workflow

This appendix series documents the analytic workflow used to reconstruct adherence-specific outcomes and develop the decision-analytic framework used in this dissertation.

The workflow proceeds in four stages:

1. **Appendix A1:** Derivation of placebo adherent and non-adherent subgroup outcomes during the uncontaminated 3-year DPP period using published adherence prevalence and relative risk estimates.
2. **Appendix A2:** Estimation of non-adherent Lifestyle and Metformin subgroup outcomes using a constant-mixture identification framework combining ITT and adherent-only outcomes.
3. **Appendix A3:** Assembly of subgroup-specific cost, QALY, and adherence inputs used in the decision-analytic model.
4. **Appendix A4:** Construction of the adherence-testing decision framework comparing Test versus No-Test treatment strategies under imperfect adherence classification.

Together, these appendices provide the methodological foundation for evaluating how adherence classification, subgroup reconstruction, and counterfactual assumptions influence cost-effectiveness conclusions in the DPP.

Appendix A1.

Analytic RoadMap: Deriving 3-Year Placebo Adherent and Non-Adherent Outcomes (QALYs and Costs):

This appendix develops a method for estimating separate 3-year costs and QALYs for adherent and non-adherent placebo participants during the uncontaminated randomized phase of the Diabetes Prevention Program (DPP) (Years 1–3).

Because the DPP did not directly report subgroup-specific outcomes for placebo adherence groups, these quantities must be inferred indirectly using published aggregate trial outcomes, adherence prevalence estimates, and relative differences in diabetes incidence between adherent and non-adherent placebo participants.

The appendix proceeds in five steps:

1. Identification of placebo adherence-specific diabetes risks,
2. Construction of subgroup-specific QALYs,
3. Estimation of subgroup-specific costs,
4. Comparison of implied subgroup differences,

5. Recalculation of incremental cost-effectiveness results using adherence-adjusted placebo comparators.

This framework provides the basis for later decision-analytic modeling and adherence-testing analyses presented in subsequent appendices.

Identification Strategy for Placebo Adherence Subgroups

The DPP did not directly report separate 3-year costs and QALYs for clinically adherent and non-adherent placebo participants. Therefore, subgroup outcomes must be inferred indirectly.

This identification strategy combines three published quantities:

1. Overall placebo diabetes incidence during the uncontaminated DPP period (Years 1–3), **(0.281)**
2. Placebo adherence prevalence reported by Walker et al., **(0.77)**
3. The relative reduction in diabetes incidence among adherent versus non-adherent placebo participants reported by Walker et al. **(0.09)**

Together, these quantities form a system of equations that uniquely identifies subgroup-specific diabetes risks for adherent and non-adherent placebo participants.

These risks are then combined with diabetes-state-specific costs and QALYs from the DPP ITT analysis to estimate expected subgroup-specific outcomes.

Table A1.1: The contaminated **APS 10-year outcomes** (undiscounted costs and QALYs, **original QALYs**, Health system perspective, under the implicit universal placebo adherence assumption used in Herman et al. (2013)

Arm	Cost ($)	QALYs
Lifestyle	28,028	6.80
Metformin	27,151	6.74
Placebo	28,236	6.67

Table A1.2: The uncontaminated APS outcomes; undiscounted costs and QALYs, **original QALYs**, Health system perspective, under the implicit universal placebo adherence assumption used in Herman et al. (2013). QALY values were recovered from Herman (2013) published histograms when tabulated values were unavailable.

Arm	Cost ($)	QALYs
Lifestyle	8,835	2.070
Metformin	6,169	1.991
Placebo	5,867	1.996

Table A1.3a: The uncontaminated adherent-only 3-year outcomes (undiscounted costs and QALYs, **altered QALYs**, (based on conventional QALY calculation method) Health system perspective, under the implicit universal placebo adherence assumption used in Herman et al. (2013)

Arm	Cost ($)	QALYs
Lifestyle	8,835	2.0715
Metformin	6,169	2.0125
Placebo	5,867	2.0060

Table A1.3b: Altered QALY calculations: Based on conventional method; using averages of beginning- and end-of-year utility values.

	Original QALYs	Altered QALYs
Baseline utility	NA	0.71
LS - TOTAL	**2.070**	**2.0715**
Year 1	0.696	0.703
Year 2	0.667	0.6815
Year 3	0.707	0.687
MET-TOTAL	**1.991**	**2.0125**
Year 1	0.677	0.6935
Year 2	0.647	0.6620
Year 3	0.667	0.6570
PBO – TOTAL	**1.996**	**2.0060**
Year 1	0.672	0.691
Year 2	0.634	0.653
Year 3	0.690	0.662

Altered QALYs calculation Steps:

Using this general formula; QALY$_j$ = Σt=1 to 3 [(U_j, t−1 + U_j, t) / 2] = $U_0/2 + U_1+U_2 + U_3/2$

Original QALYs were altered as follows:

Lifestyle Intervention (LS)
$Year_1$ = 0.71/2 + 0.696/2 = 0.355 + 0.348 = **0.703**
$Year_2$ = 0.696/2 + 0.667/2 = 0.348+ 0.3335 = **0.6815**
$Year_3$ = 0.667/2 + 0.707/2 = 0.3335 + 0.3535 = **0.687**

Metformin
$Year_1$ = 0.71/2 + 0.677/2 = 0.355 + 0.3385 = **0.6935**
$Year_2$ = 0.677/2 + 0.647/2 = 0.3385+ 0.3235 = **0.6620**
$Year_3$ = 0.647/2 + 0.667/2 = 0.3235 + 0.3335 = **0.6570**

PBO
$Year_1$ = 0.71/2 + 0.672/2 = 0.355 + 0.336 = **0.691**
$Year_2$ = 0.672/2 + 0.634/2 = 0.336+ 0.317 = **0.653**
$Year_3$ = 0.634/2 + 0.690/2 = 0.317 + 0.345 = **0.662**

Table A1.4: 3-Year Cumulative Diabetes Incidence: DPP clinical trial results reported by the Diabetes Prevention Program Research Group (2002).[2]

Converting incidence rates per 100 PY to cumulative 3-year risk using:

Cumulative 3-year Incidence **= $1-e^{-rt}$** (assuming a constant continuous incidence rate)

Arm	Incidence rate per 100 PY	Incidence rate per 1 PY	Cumulative 3-year Diabetes Incidence (%)
LS	4.8	0.048	0.134
MET	7.8	0.078	0.209
PBO	11.0	0.110	0.281

Step (1): **Identification of placebo adherence-specific diabetes risks**

This step identifies the 3-year diabetes risks for adherent and non-adherent placebo participants. These risks are not directly observed in the DPP but can be recovered from three known quantities:

1. The **Cumulative 3-year placebo diabetes incidence** (Table A1.4),
2. The **placebo adherence prevalence (77%)** from Walker et al. (2006), and
3. The **relative risk reduction (RRR)** for adherent versus non-adherent placebo participants (9%); RR of 0.91 from Walker et al. (2006).

Let $\mathbf{P(D)_{ADH}}$ denote the 3-year probability(risk) of developing diabetes among adherent placebo participants, and

$\mathbf{P(D)_{nADH}}$ denote the corresponding probability (risk) among non-adherent placebo participants.

The Walker et al. relative risk reduction for adherent versus non-adherent placebo participants (9%) implies the following relationship between the two subgroup risks:

$$P(D)_{ADH} = 0.91 \times P(D)_{nADH} \qquad \text{Equation (1)}$$

Conceptually, the observed placebo incidence (0.281) can be viewed as a weighted average of diabetes risk among adherent and non-adherent placebo participants. The Walker et al. relative risk estimate links the two subgroup risks, allowing the system to be solved algebraically.

At the same time, the observed overall placebo diabetes incidence over 3 years (0.281) must equal the weighted average of the two subgroup risks:

$$0.281 = \pi \times P(D)_{ADH} + (1 - \pi) \times P(D)_{nADH} \qquad \text{Equation (2)}$$

where $\pi =$ adherence prevalence (0.77) in the placebo arm and, $\pi - 1$(0.23) is the non-adherence prevalence.

Equations (1) and (2) form a system of two equations in two unknowns. Solving this system yields unique values for $P(D)_{ADH}$ and $P(D)_{nADH}$, as shown below:

Substitute Equation (1) and the given variables into Equation (2)

0.281 = 0.77 * (0.91P(D)nADH) +0.23 * P(D)nADH Equation (3)

0.281 = P(D)nADH * (0.77 * 0.91+ 0.23*1) Equation (4)

0.281 = 0.9307 * P(D)nADH Equation (5)

$P(D)_{nADH} = \frac{0.281}{0.9307} \approx \mathbf{0.302}$ Equation (6)

Substitute Equation (6) into Equation (1)

$P(D)_{ADH} = 0.91 * 0.302 \approx \mathbf{0.275}$ Equation (7)

Step (2): Expected-value construction of placebo adherence-specific QALYs

Having identified subgroup-specific diabetes risks in Step (1), we now construct the corresponding 3-year QALYs for adherent and non-adherent placebo participants.

We assume that, over the 3-year uncontaminated horizon, health-related quality of life depends only on whether a patient develops diabetes. Under this assumption, each individual belongs to one of two health states:

- Diabetes (D), or
- No diabetes (ND),

each with an associated 3-year QALY total obtained from the (2012) DPP ITT cost effectiveness analysis supplementary data (Table A3), reported here in table A1.5 below:

Expected QALYs for each adherence subgroup are therefore computed as weighted averages of these two state-specific QALYs, using the diabetes risks identified in Step (1).

Table A1.5: Diabetic and Non-Diabetic PBO QALYs from the (2012) DPP Research group ITT cost effectiveness analysis supplementary data (Table 3); Original and Altered QALYs.

	Original	**Altered QALYs**
Baseline Utility	**NA**	**0.71**
Diabetic QALYs (D) -Total	**2.032**	**2.0555**
Year 1	0.688	0.699
Year 2	0.681	0.6845
Year 3	0.663	0.672
Nondiabetic QALYs (ND) - Total	**2.044**	**2.0600**
Year 1	0.689	0.6995
Year 2	0.677	0.683
Year 3	0.678	0.6775

Altered QALY calculations: Based on conventional method; using averages of beginning- and end-of-year utility values.

Using the general formula: $\mathbf{QALY_j} = \sum^3 (U^j (t\text{-}1) + U^j / 2 = U_0/2 + U_1 + U_2 + U_3/2.$

Diabetes status QALYs were altered as follows:

Diabetic QALYs

$Year_1 = 0.71/2 + 0.688/2 = 0.355 + 0.344 = \mathbf{0.699}$

$Year_2 = 0.688/2 + 0.681/2 = 0.344 + 0.3405 = \mathbf{0.6845}$

$Year_3 = 0.681/2 + 0.663/2 = 0.3405 + 0.3315 = \mathbf{0.672}$

Non-Diabetic QALYs
$Year_1$= 0.71/2 + 0.689/2 = 0.355+0.3445 = **0.6995**
$Year_2$ = 0.689/2 + 0.677/2 = 0.3445+0.3385 = **0.683**
$Year_3$ = 0.677/2+0.678/2 = 0.3385 + 0.339 = **0.6775**

From Equation (6) and (7):

$P(D)_{ADH}$ = **0.275**

$P(D)_{nADH}$ = **0.302**

So:

$P(ND)_{ADH}$ = 1- 0.275 = **0.725**

$P(ND)_{nADH}$ = 1 - 0.302 = **0.698**

Using the weighted average equation:

$Q_{ADH} = P(ND)_{ADH} * Q_3^{ND} + P(D)_{ADH} * Q_3^{D}$ Equation (8)

Q_{ADH} = 0.725 * 2.0600 + 0.275 * 2.0555 = **2.0588** Equation (9)

$Q_{nADH} = P(ND)_{nADH} * Q_3^{ND} + P(D)_{nADH} * Q_3^{D}$ Equation (10)

Q_{nADH} = 0.698 * 2.0600 + 0.302 * 2.0555 = **2.0586** Equation (11)

Step (3): QALY Difference = 2.0588 - 2.0586 **= 0.0002/QALYs over 3 years**

Step (4): **Expected-value construction of placebo adherence-specific costs**

Analogously, we now compute expected 3-year medical costs for adherent and non-adherent placebo participants.

We assume that, over the 3-year uncontaminated period, medical costs depend only on diabetes status (Diabetes vs No Diabetes). The 2012 DPP Research group ITT Cost-effectiveness analysis supplementary data Table 3 (reported here in Table A1.6a) provide the corresponding 3-year outside program costs for these two states, and the Table 1 of the main article[1] (reported here in Table A1.6b) provide the 3-year inside-program costs.

Expected subgroup-specific costs are therefore obtained by weighting diabetes-state costs by the subgroup-specific diabetes risks identified in Step (2), total inside program cost is then added.

Given the following inputs:
Table A1.6a: PBO outside-DPP medical costs by diabetes status over the uncontaminated 3-year DPP period (ITT):

Diabetes Status	Year 1	Year 2	Year 3	3-year Total
Diabetes (D)	$2,705	$3,724	$3,638	**$10,067**
No diabetes (ND)	$1,978	$2,106	$2,431	**$6,515**

Note: Outside-DPP medical costs represent diabetes-related and non–diabetes-related medical expenditures incurred outside the intervention program. Values are taken from the Diabetes Prevention Program Research Group (2012) cost-effectiveness analysis based on intention-to-treat (ITT) estimates during Years 1–3.

Table A1.6b: Inside-DPP medical costs over the uncontaminated 3-year DPP period (ITT):

Arm	Year 1	Year 2	Year 3	3-year Total
Lifestyle	1,826	887	915	**3,628**
Metformin	584	294	299	**1,177**
Placebo	87	50	47	**184**

Note: Inside-DPP costs reflect intervention delivery and program costs incurred regardless of diabetes status. Values are taken from the Diabetes Prevention Program Research Group cost-effectiveness analysis based on intention-to-treat (ITT) estimates during Years 1–3.

From Equation (6) and (7):

$P(D)_{ADH}$ = **0.275**

$P(D)_{nADH}$ = **0.302**

So:

$P(ND)_{ADH}$ = 1- 0.275 = **0.725**

$P(ND)_{nADH}$ = 1 - 0.302 = **0.698**

Using the weighted average equation:

C_{ADH} = Inside Cost + $P(ND)_{ADH} * C_3^{ND} + P(D)_{ADH} * C_3^{D}$ Equation (12)
C_{ADH} = $184 + 0.725 * 6,515 + 0.275 * 10,067 = **$7,676** Equation (13)

C_{nADH} = Inside Cost + $P(ND)_{nADH} * C_3^{ND} + P(D)_{nADH} * C_3^{D}$ Equation (14)
C_{nADH} =$184+ 0.698* 6,515 + 0.302 * 10,067 = **$7,772** Equation (15)

Step (5): The implied cost difference = $7,772 – 7,676 **= $96 over 3years**

Table A1.7: Final 3-Year Placebo Subgroup Outcomes (uncontaminated):

Subgroup	Cost ($)	QALYs
Placebo ADH	7,676	2.0588
Placebo nADH	7,772	2.0586

Interpretation:
The reconstructed placebo adherence subgroups produce only small QALY differences over the 3-year horizon, but meaningful cost differences driven by differential diabetes incidence. These reconstructed subgroup outcomes are later used as inputs in the adherence-testing decision framework.

Table A1.8: Original ICERs with **original QALYs;** deterministic analysis; healthcare perspective, undiscounted cost and QALYs, contaminated 10-year DPP horizon, under the implicit universal placebo adherence assumption used in Herman et al. (2013)

Outcome	**Lifestyle vs Placebo**	**Metformin vs Placebo**	**Lifestyle vs Metformin**
Δ Costs ($)	-210	-1,086	877
Δ QALYs	0.14	0.08	0.06
ICER ($/QALY)	Cost-saving	Cost-saving	14,213* (**14,617**)

*Reported by Herman

Table A1.9
Appropriate ICERs with interventions ordered by increasing effectiveness; under the implicit universal placebo adherence assumption used in Herman et al. (2013) **altered QALYs,** health care perspective, deterministic analysis, uncontaminated 3-year DPP horizon.

Intervention (1)	QALYs (2)	Costs (3)	ICERs vs NME (4)	NME (5)	ICERs vs NME2 (6)	NME 2 (7)
PBO	2.0060	$5,867		NA		NA
MET	2.0125	$6,169	$46,462	PBO		
LS	2.0715	$8,835	$45,186	MET	$45,313	PBO

ICERs is incremental cost-effectiveness ratio; **LS** is the lifestyle modification intervention, **MET** is metformin, **NA** is not applicable, **NME** is next most effective intervention not dominated; **NME2** is next most effective intervention not extendedly dominated; **PBO** is placebo; **QALY**s are quality adjusted life years.
Note: The placebo comparator in this table corresponds to the adherent placebo subgroup, which represents the reference outcomes reported by Herman (2013).

Table A1.10
Appropriate ICERs under the implicit universal placebo adherence assumption used in Herman et al. (2013), **original QALYs,** health care perspective, uncontaminated 3-year DPP horizon.

Intervention (1)	QALYs (2)	Costs (3)	ICERs vs NME (4)	NME (5)	ICERs vs NME2 (6)	NME 2 (7)
MET	1.991	$6,169	DOM by PBO	NA		NA
PBO	1.996	$5,867		PBO		NA
LS	2.070	$8,835	$40,108	PBO		NA

Table A1.11: True PBO ADH
Appropriate ICERs; **PBO** adherence adjustment (PBO ADH QALYs and Cost); **Altered QALYs,** health care perspective, uncontaminated 3-year DPP horizon.

Intervention (1)	QALYs (2)	Costs (3)	ICERs vs NME (4)	NME (5)	ICERs vs NME2 (6)	NME 2 (7)
MET	2.0125	$6,169		NA		NA
PBO ADH	2.0588	$7,676	$32,549	MET		NA
LS	2.0715	$8,835	$91,260	PBO(ADH)		MET

Table A1.12: True PBO nADH
Appropriate ICERs; **PBO** adherence adjustment (PBO nADH QALYs and Cost); **Altered QALYs,** health care perspective, uncontaminated 3-year DPP horizon.

Intervention (1)	QALYs (2)	Costs (3)	ICERs vs NME (4)	NME (5)	ICERs vs NME2 (6)	NME 2 (7)
MET	2.0125	$6,169		NA		NA
PBO nADH	2.0586	$7,772	$34,772	MET		NA
LS	2.0715	$8,835	$82,403	PBO (nADH)		MET

Appendix A2: **Estimating Non-Adherent Lifestyle and Metformin Outcomes Using Mixture Decomposition**

Unlike the placebo arm, the DPP provides both intention-to-treat (ITT) outcomes and adherent-only outcomes for the Lifestyle (LS) and Metformin (MET) interventions.

However, subgroup-specific outcomes for non-adherent participants were not directly reported.

To recover these missing subgroup outcomes, this appendix applies a constant-mixture identification framework in which the ITT outcome is treated as a weighted average of adherent and non-adherent subgroup outcomes.

Under this framework, observed ITT outcomes, adherence prevalence, and adherent-only outcomes jointly identify the implied non-adherent subgroup outcomes.

This procedure is applied separately to costs and QALYs for both LS and MET.

General mixture formula:

For any arm (Lifestyle or Metformin), we assume:

$$O^{ITT} = p(ADH) \cdot O^{ADH} + \left(1 - p(ADH)\right) \cdot O^{nADH}$$

Where:

$\mathbf{O^{ITT}}$ denotes the mean outcome observed under the intention-to-treat (ITT) analysis (e.g., cost or QALYs averaged over all participants assigned to a treatment arm).
$\mathbf{p(ADH)}$ is the proportion of participants who were adherent in that treatment arm.
$\mathbf{O^{ADH}}$ is the mean outcome among adherent participants, and
$\mathbf{O^{nADH}}$ is the mean outcome among non-adherent participants.

The mixture relationship: $O^{ITT} = p(ADH) \cdot O^{ADH} + \left(1 - p(ADH)\right) \cdot O^{nADH}$
follows from the ITT mean being a weighted average of adherent and non-adherent outcomes. Rearranging yields the implied non-adherent outcome:

$$O^{nADH} = \frac{O^{ITT} - p(ADH) \cdot O^{ADH}}{1 - p(ADH)}$$ Equation 16

We apply this separately for **cost** and **QALYs**.

Similarly for cost; $$C^{nADH} = \frac{C^{ITT} - p(ADH) \cdot C^{ADH}}{1 - p(ADH)}$$ Equation 17

Table A2.1: Herman APS 2013: 3-years outcomes; undiscounted costs and QALYs, **altered QALYs**. QALY values were recovered from Herman (2013) reversed publication supplemental Table 3

Arm	Cost ($)	QALYs
Lifestyle	8,811	2.099
Metformin	6,386	2.0675

(1) Lifestyle non-adherent

Inputs (QALYs are altered, Herman APS QALYs are Histogram AI extracted, cost includes: inside and outside cost)

- 10-year LS Adherence prevalence (Herman 2012 and Herman 2013p4): $p(ADH) = 0.54$
- 3-year LS Adherence prevalence = $0.54^{3/10}$ = 0.8312
- 3-year ITT LS outcomes (from Herman 2012): Cost = $8,811. QALYs = 2.099
- 3-year APS LS outcomes (from Herman 2013 APS analysis): Costs = $8,835, QALYs = 2.0715

Calculations:

Substitute these values into equations 16 and 17:

- $Q^{nADH} = \frac{Q^{ITT} - p(ADH) \cdot Q^{ADH}}{1 - p(ADH)} = \frac{2.099 - 0.8312 * 2.0715}{1 - 0.8312} = 2.2344$

- $C^{nADH} = \frac{C^{ITT} - p(ADH) \cdot C^{ADH}}{1 - p(ADH)} = \frac{8{,}811 - 0.8312 * 8{,}835}{1 - 0.8312} = \$8{,}693$

(2) Metformin non-adherent QALYs (using ITT Q = 2.0585)

Inputs (QALYs are altered, Adherent-only QALYs are Histogram AI extracted, cost include inside + outside cost)

- Adherence prevalence (Walker 2006):
  $p(ADH) = 0.71$
- 3-year ITT MET outcomes (from Herman 2012,10-year ITT analysis):
  Year 1-3 Total Costs = 6,386
  Year 1-3 Total QALYs = 2.0675
- 3-year APS MET outcomes: (from Herman 2013 APS analysis)
  Costs = 6,169
  QALYs = 2.0125

Calculations:

Substituting these values into equations 16 and 17:

- $C^{nADH} = \frac{C^{ITT} - p(ADH) \cdot C^{ADH}}{1 - p(ADH)} = \frac{6{,}386 - 0.71 * 6{,}169}{1 - 0.71} = \$6{,}917$

- $Q^{nADH} = \frac{Q^{ITT} - p(ADH) \cdot Q^{ADH}}{1 - p(ADH)} = \frac{2.0675 - 0.71 * 2.0125}{1 - 0.71} = 2.2022$

Modeling Limitation: Implied Non-Adherent Subgroup Outcomes

Non-adherent subgroup outcomes for the Lifestyle (LS) and Metformin (MET) interventions were not directly observed in the DPP data and were therefore inferred indirectly using a constant-mixture identification framework. Under this approach, the observed intention-to-treat (ITT) outcome is treated as a weighted average of adherent and non-adherent subgroup outcomes, conditional on the observed adherence prevalence.

Because the adherent-only subgroup outcomes were reconstructed from separate analyses and the ITT outcomes reflect averages across both adherent and non-adherent participants, the implied non-adherent subgroup values are algebraically determined by the mixture equations rather than directly estimated from observed patient-level data. Consequently, when the observed ITT mean exceeds the reconstructed adherent-only mean, the implied non-adherent subgroup outcome may mathematically exceed the adherent subgroup outcome.

In the present application, this occurs for QALYs in both the Lifestyle and Metformin arms, producing non-adherent subgroup QALY estimates that are higher than the corresponding adherent subgroup estimates. These values should not be interpreted as evidence that non-adherent patients experience superior clinical outcomes relative to adherent patients. Rather, they represent model-implied quantities required to satisfy the internal consistency constraints of the mixture decomposition under the available published data.

Accordingly, these reconstructed subgroup estimates should be interpreted as identification-based inputs used to operationalize the decision-analytic framework, rather than as directly observed or causally interpretable clinical subgroup outcomes. The purpose of the framework is therefore methodological and structural — namely, to demonstrate how adherence classification, subgroup reconstruction, and imperfect adherence prediction influence treatment allocation and cost-effectiveness conclusions under incomplete subgroup observability.

Sensitivity analyses should therefore be interpreted with appropriate caution, particularly where subgroup outcomes are partially identified or inferred from aggregate trial summaries rather than directly measured participant-level data.

Appendix A3: Decision Tree Model Inputs

This appendix summarizes subgroup-specific 3-year cost and QALY inputs for **Lifestyle, Metformin, and Placebo**, each stratified by adherent and non-adherent status. This appendix also provides the Diagnostic Classification Parameters used for Adherence Testing, as well as Parameter Ranges Used in Sensitivity Analyses.

Table A3.1: Model Inputs for Cost and QALY Outcomes (3-Year Horizon)

Intervention	Group	3-Year Cost ($)	3-Year QALYs	Adherence Prevalence	Source
Lifestyle	Adherent	8,835	2.0715	10-Yr adherence Prevalence = 0.54; converted 3-year prevalence ($0.54^{3/10}$ = 0.8312)	Herman (2013)
Lifestyle	Non-Adherent	8,693	2.2344*	0.46(1-0.8312=0.1688)	
Metformin	Adherent	6,169	2.0125	0.71	Walker (2006)
Metformin	Non-Adherent	6,917	2.2022*	0.29	
Placebo	Adherent	7,676	2.0588	0.77	Walker (2006)
Placebo	Non-Adherent	7,772	2.0586	0.23	

* **Modeling limitation:** Non-adherent subgroup outcomes for LS and MET were inferred using the mixture decomposition framework described in Appendix A2. In some cases, implied non-adherent QALYs exceed adherent QALYs because subgroup values are algebraically identified from aggregate ITT and adherent-only outcomes rather than directly observed clinical data. These values should therefore be interpreted as model-implied quantities rather than causal subgroup estimates.

Table A3.2

Diagnostic Classification Parameters for Adherence Testing

Parameter	Base-Case Value
Sensitivity	0.90
Specificity	0.90

Table A3. 3

Parameter Ranges Used in Sensitivity Analyses

Parameter	Base Case	Range Evaluated	Source / Rationale
LS adherence prevalence	10-Yr adherence Prevalence = 0.54; converted 3-year prevalence ($0.54^{3/10}$ = 0.8312	0.40 – 0.70	Herman et al., 2013; plausible variation
MET adherence prevalence	0.71	0.55 – 0.85	Walker et al., 2006
PBO adherence prevalence	0.77	0.60 – 0.90	Walker et al., 2006
Sensitivity of adherence classification	0.90	0.70 – 0.95	Assumption (base-case 0.9)
Specificity of adherence classification	0.90	0.70 – 0.95	Assumption (base-case 0.9)
Willingness-to-pay threshold	$50,000	$50,000–$100,000	Conventional thresholds
Time horizon	3 years	—	Uncontaminated randomized DPP period

Appendix A4: **Decision-Analytic Framework and Identification**

This appendix develops a decision-analytic framework that evaluates whether adherence classification testing can improve treatment selection during the uncontaminated 3-year DPP period.

The framework assumes that patients have an underlying but unobserved adherence type (adherent or non-adherent). Because adherence status is unknown at treatment initiation, treatment decisions must instead rely on an imperfect adherence-classification test characterized by sensitivity and specificity.

The model evaluates whether using this imperfect classification system to guide treatment allocation improves expected costs and QALYs relative to a conventional "No-Test" strategy in which all patients receive the same intervention.

The framework consists of two linked components:

1. A classification layer that determines predicted adherence status,
2. An outcome layer that assigns expected costs and QALYs conditional on the predicted classification.

Classification Layer

The adherence test generates predicted adherence status:

- Test positive (designated adherent); with marginal probability; **p (“ADH”)**
- Test negative (designated non-adherent); with marginal probability **p(“nADH”)**

These probabilities depend on:

- adherence prevalence (π)
- test sensitivity (Se)
- test specificity (Sp)

Together these parameters determine the probabilities of true positives, false positives, true negatives, and false negatives**. These are joint probabilities:**

TP = p (“ADH”, ADH)

FP = p (“ADH”, nADH)

TN = p (“nADH”, nADH)

FN = p (“nADH”, ADH)

Outcome Layer

The outcome layer uses subgroup-specific 3-year costs and QALYs for adherent and non-adherent patients under three interventions:

- Lifestyle (LS)
- Metformin (MET)
- Placebo (PBO)

For each test result (positive or negative), expected costs and QALYs are calculated for each intervention by combining the conditional probability of true adherence given the test result (PPV or NPV),

PPV = P (ADH | “ADH”)

NPV = P (nADH | “nADH”)

And the corresponding subgroup-specific outcomes for adherent and non-adherent patients.

For each test result, the intervention with the highest expected net monetary benefit (NMB) is selected.

Comparison Strategy

Expected outcomes under the Test strategy are compared with a No-Test strategy**.**

Under the No-Test strategy, all patients receive Lifestyle**,** consistent with the original intention-to-treat cost-effectiveness result in which Lifestyle was the preferred intervention.

Inputs:

Prevalence & test performance

- True adherence prevalence (π): **0.54**
- Sensitivity (Se): **0.90**
- Specificity (Sp): **0.90**

Subgroup Outcomes (uncontaminated; 3-year horizon)

Lifestyle (LS)

- Adherent: $C^{ADH} = \$8{,}835; \quad Q^{ADH} = 2.0715$
- Non-Adherent: $C^{nADH} = \$8{,}693; \; Q^{nADH} = 2.2344$

Placebo (PBO)

- Adherent: $C^{ADH} = \$7{,}704; \quad Q^{ADH} = 2.0587$
- Non-Adherent: $C^{nADH} = \$7{,}804; \; Q^{nADH} 2.0586$

Metformin (MET)

- Adherent: $C^{ADH} = \$6{,}169; \quad Q^{ADH} = 2.0125$
- Non-Adherent: $C^{nADH} = \$6{,}917; \; Q^{nADH} 2.2022$

No Test comparator:

Under the **No-Test strategy**, all patients receive treatment based on the conventional intention-to-treat (ITT) cost-effectiveness result.

Using the 3-year ITT outcomes for Lifestyle:

- $C_{NoTest} = \$8{,}811$
- $Q_{NoTest} = 2.099$ (Altered QALYs)

Treatment rule under the Test strategy:
For each test result (ADH and nADH), expected costs and QALYs are calculated for LS, MET, and PBO using the conditional adherence probabilities implied by the test result. Expected NMB is then calculated for each intervention, and the intervention with the highest expected NMB is selected.

Calculations:

1. Step 1 – Marginal test result probabilities: P("ADH") and P("nADH")

Using:

$$p(\mathrm{ADH}) = \mathrm{Se} \cdot \pi + (1 - \mathrm{Sp}) \cdot (1 - \pi)$$
$$p(\text{"nADH"}) = (1 - \mathrm{Se}) \cdot \pi + \mathrm{Sp} \cdot (1 - \pi)$$

With $\pi = 0.54$, Se = Sp = 0.90:

- $p(\mathrm{ADH}) = 0.9 \times 0.54 + 0.1 \times 0.46 = 0.486 + 0.046 = 0.532$
- $p(\text{"nADH"}) = 1 - p(\mathrm{ADH}) = 1 - 0.532 = 0.468$

So:

- About **53.2%** of patients test "ADH" (positive)
- About **46.8%** of patients test "nADH" (negative)

2. Step 2 – PPV and NPV (conditional probabilities)

2.1 Positive test ("ADH result"): PPV

$$\mathrm{PPV} = P(\,\mathrm{ADH} \mid \text{"ADH"}\,) = \frac{\mathrm{Se}\cdot\pi}{p(\text{"ADH"})} = \frac{0.9\times0.54}{0.532} = \frac{0.486}{0.532} \approx 0.914 \qquad \text{Eqn (1)}$$

So, among **test-positive ("ADH")** patients:

- **91.4%** are truly adherent
- **8.6%** are truly non-adherent

2.2 Negative test ("nADH result"): NPV

Firstly, we compute $P\big(\text{ADH} \mid p(\text{"nADH"})\big)$: false negatives among those predicted non-adherent

$$\left(P\big(\text{ADH} \mid p(\text{"nADH"})\big) = \frac{(1-\text{Se})\cdot\pi}{(1-\text{Se})\cdot\pi+\text{Sp}\cdot(1-\pi)} = \frac{0.1\times0.54}{0.1\times0.54+0.9\times0.46}\right. \qquad \text{Eqn (2)}$$

$$= \frac{0.054}{0.054+0.414} \approx 0.115$$

Then:

$$\text{NPV} = P\big(\text{nADH} \mid p(\text{"nADH"})\big) = 1 - P\big(\text{ADH} \mid (\text{"nADH"})\big) \approx 1 - 0.115 = 0.885$$

So, among **test-negative ("nADH")** patients:

- **88.5%** are truly non-adherent
- **11.5%** are actually adherent (False negative)

3. Step 3 – Effective outcomes conditional on the test result

3.1: Patients with **positive test**:

Among ADH patients:

- Proportion truly adherent = $P(\text{ADH} \mid \text{"ADH"}) = 0.914$
- Proportion truly non-adherent = $1 - P(\text{ADH} \mid \text{"ADH"}) = 0.086$

So, **LS outcomes given ADH test result** are:

$$C_{LS} = P(\text{ADH} \mid \text{"ADH"}) \cdot C_{LS}^{ADH} + P(\text{nADH} \mid \text{"ADH"}) \cdot C_{LS}^{nADH}$$

$$Q_{LS} = P(\text{ADH} \mid \text{"ADH"}) \cdot Q_{LS}^{ADH} + P(\text{nADH} \mid \text{"ADH"}) \cdot Q_{LS}^{nADH}$$

- $C_{LS} = 0.914 \times 8{,}835 + 0.086 \times 8693 = \$8{,}823$
- $Q_{LS} = 0.914 \times 2.0715 + 0.086 \times 2.2344 = 2.0855$

MET outcomes given ADH test result are:

$$C_{MET} = P(ADH \mid "ADH") \cdot C_{MET}^{ADH} + P(nADH \mid "ADH") \cdot C_{MET}^{nADH}$$

$$Q_{MET} = P(ADH \mid "ADH") \cdot Q_{MET}^{ADH} + P(nADH \mid "ADH") \cdot Q_{MET}^{nADH}$$

- $C_{MET} = 0.914 \times 6{,}169 + 0.086 \times 6{,}917 = \$6{,}233$
- $Q_{MET} = 0.914 \times 2.0125 + 0.086 \times 2.2022 = 2.029$

PBO outcomes given ADH test result are:

$$C_{PBO} = P(ADH \mid "ADH") \cdot C_{PBO}^{ADH} + P(nADH \mid "ADH") \cdot C_{PBO}^{nADH}$$
$$Q_{PBO} = P(ADH \mid "ADH") \cdot Q_{PBO}^{ADH} + P(nADH \mid "ADH") \cdot Q_{PBO}^{nADH}$$

- $C_{PBO} = 0.914 \times 7{,}704 + 0.086 \times 7{,}804 = \$7{,}712$
- $Q_{PBO} = 0.914 \times 2.0587 + 0.086 \times 2.0586 = 2.0587$

3.2: Patients with **negative test**

Among nADH patient:

- Proportion truly non-adherent = NPV = 0.885
- Proportion truly adherent = 1 − NPV = 0.115

So, **LS outcomes given nADH** are:

$$C_{LS} = P(ADH \mid "nADH") \cdot C_{LS}^{ADH} + P(nADH \mid "nADH") \cdot C_{LS}^{nADH}$$
$$Q_{LS} = P(ADH \mid "nADH") \cdot Q_{LS}^{ADH} + P(nADH \mid "nADH") \cdot Q_{LS}^{nADH}$$

- $C_{LS} = 0.115 \times 8{,}835 + 0.885 \times 8{,}693 = \$\mathbf{8{,}709}$
- $Q_{LS} = 0.115 \times 2.0715 + 0.885 \times 2.2344 = \mathbf{2.2157}$

MET outcomes given nADH test result are:

$$C_{MET} = P(ADH \mid "nADH") \cdot C_{MET}^{ADH} + P(nADH \mid "nADH") \cdot C_{MET}^{nADH}$$
$$Q_{MET} = P(ADH \mid "nADH") \cdot Q_{MET}^{ADH} + P(nADH \mid "nADH") \cdot Q_{MET}^{nADH}$$

- $C_{MET} = 0.115 \times 6{,}169 + 0.885 \times 6{,}917 = \$\mathbf{6{,}831}$
- $Q_{MET} = 0.115 \times 2.0125 + 0.885 \times 2.2022 = \mathbf{2.1804}$

PBO outcomes given nADH test result are:

$$\mathrm{C_{PBO}} = \mathrm{P(ADH \mid "nADH")} \cdot \mathrm{C_{PBO}^{ADH}} + \mathrm{P(nADH \mid "nADH")} \cdot \mathrm{C_{PBO}^{nADH}}$$
$$\mathrm{Q_{PBO}} = \mathrm{P(ADH \mid "nADH")} \cdot \mathrm{Q_{PBO}^{ADH}} + \mathrm{P(nADH \mid "nADH")} \cdot \mathrm{Q_{PBO}^{nADH}}$$

- $\mathrm{C_{PBO}} = 0.115 \times 7{,}704 + 0.885 \times 7{,}804 = \$\mathbf{7,793}$
- $\mathrm{Q_{PBO}} = 0.115 \times 2.0587 + 0.885 \times 2.0586 = \mathbf{2.0586}$

4. Step 4 – Treatment selection under the Test strategy

For each test result, expected net monetary benefit (NMB) is calculated for LS, MET, and PBO:

$$\mathrm{NMB} = \lambda \mathrm{Q} - \mathrm{C}$$

with base-case willingness-to-pay threshold

$$\lambda = \frac{\$50{,}000}{\mathrm{QALY}}$$

4.1 Positive test result (ADH):

Using the expected outcomes calculated in step 3:

- LS: $\mathrm{C_{LS}} = 8{,}823,\ \mathrm{Q_{LS}} = 2.0855$ Eqn (3)
- MET: $\mathrm{C_{MET}} = 6{,}233,\ \mathrm{Q_{MET}} = 2.029$
- PBO: $\mathrm{C_{PBO}} = 7{,}712,\ \mathrm{Q_{PBO}} = 2.0587$

Expected NMBs are:

$$NMB_{LS} = 50{,}000(2.0855) - 8{,}823 = \mathbf{95,452}$$
$$NMB_{MET} = 50{,}000(2.029) - 6{,}233 = 95{,}217$$
$$NMB_{PBO} = 50{,}000(2.0587) - 7{,}712 = 95{,}223$$

Therefore, for patients with a positive adherence test result: Choose LS

4.2 Negative test result (nADH)

Using the expected outcomes calculated in step 3:

- LS: $C_{LS} = 8{,}709,\ Q_{LS} = 2.2157$
- MET: $C_{MET} = 6{,}831,\ Q_{MET} = 2.1804$ Eqn (4)
- PBO: $C_{PBO} = 7{,}793,\ Q_{PBO} = 2.0586$

Expected NMBs are:

$$\text{NMB}_{\text{LS}} = 50{,}000(2.2157) - 8{,}709 = 102{,}076$$
$$\text{NMB}_{\text{MET}} = 50{,}000(2.1804) - 6{,}831 = \mathbf{102,189}$$
$$\text{NMB}_{\text{PBO}} = 50{,}000(2.0586) - 7{,}793 = 95{,}137$$

Therefore, for patients with a negative adherence test result: Choose MET, because MET has the highest expected NMB.

5. Step 5 – Overall expected outcomes for the Test strategy

From Step 1: $P(\text{"ADH"}) = 0.532,$ $P(\text{"nADH"}) = 0.468$

From Step 4, the optimal treatment rule is:

- $Test\ ADH \quad \rightarrow LS$
- $Test\ nADH \quad \rightarrow MET$

Therefore: $$\text{C}_{\text{Test}} = \text{P("ADH")}\,.\ \text{C}_{\text{LS}} + \text{P("nADH")}\,.\text{C}_{\text{MET}}$$

$$\text{Q}_{\text{Test}} = \text{P("ADH")}\,.\text{Q}_{\text{LS}} + \text{P("nADH")}\,.\text{Q}_{\text{MET}}$$

Substituting cost and QALY values for LS and MET from eqn (3) and (4):

$$\text{C}_{\text{Test}} = 0.532(8{,}823) + 0.468(6{,}831) = 6{,}512.864$$
$$\text{Q}_{\text{Test}} = 0.532(2.0855) + 0.468(2.1804) = 2.0998552$$

So, the expected outcomes for the Test strategy are:

C_{Test} = $6,513

Q_{Test} = 2.0999

1. **Step 6 – Head-to-head Incremental comparison: Test vs No Test:**

Using the 3-year ITT as no-Test Comparator outcomes, (where LS was the cost-effective intervention):

- C_{NoTest} = $8,811
- Q_{NoTest} = 2.099 (Altered QALYs)

Incremental differences for No Test relative to Test are:

$$\Delta C = C_{NoTest} - C_{Test} = 8{,}811 - 6{,}512.864 = \mathbf{2,298.136}$$
$$\Delta Q = Q_{NoTest} - Q_{Test} = 2.099 - 2.0998552 = \mathbf{-0.0008552}$$

Therefore:

- No Test costs about **$2,298 more**
- No Test yields about **0.00086 fewer QALYs**

Accordingly, the **Test strategy** dominates **No Test** in this specification: **Test** strategy is both **less costly** and **slightly more effective**.

At a WTP of $50,000/QALY, the corresponding NMBs are:

$$\text{NMB}_{\text{Test}} = 50{,}000(2.0998552) - 6{,}512.864 = \mathbf{98{,}479.896}$$

$$\text{NMB}_{\text{NoTest}} = 50{,}000(2.099) - 8{,}811 = 96{,}139$$

Therefore, the Test strategy has the higher expected NMB and is preferred at the base-case threshold.