



Evaluation protocol, version 3

LINDBERGH FORMATION — 16 rue de Maillé, 91310 Montlhéry, France · Companies register (RCS) Evry 817 946 114

Document issued on 24/09/2026

Written and versioned before the campaign was run · fictional document

Fictional document — demonstration example

This document is produced by the Blue Lemon Agent website to illustrate what an AI agent hands to its users. The companies, matters, amounts and people in it are invented. It has no legal, accounting or commercial value.

Fictional document, produced for the demonstration. The protocol, datasets and thresholds are invented.

The dataset used, and what authorises it

Element	Value	What authorises it
Claim files	640, drawn from the last twelve months	Internal control purpose, dataset pseudonymised before extraction
Pseudonymisation	Member and care provider identifiers replaced	Verified before every campaign, control recorded in the framework
Dataset authorisation period	Six months, dated purge at expiry	Register of evaluation datasets
Place of execution	Isolated resource hosted in France	The agent's architecture

The four families and their thresholds, set before execution

LINDBERGH FORMATION — a simplified joint-stock company with share capital of €250,000 · Registered office: 16 rue de Maillé, 91310 Montlhéry, France · Companies register (RCS) Evry 817 946 114 · SIRET 817 946 114 00029 · VAT FR 60 817 946 114 · "Blue Lemon Agent" trade mark · Training-provider registration no. 119 108 635 91 (Île-de-France prefect — this registration does not amount to State approval) · Qualiopi certified, no. 03896, for the "training actions" category · contact@bluelemonagent.ai · dpo@bluelemonagent.ai · Hosted in France.

Family	What it tests	Acceptance threshold	Number of cases
Routine treatments	Performance on ordinary volume	$\geq 92\%$	380
Atypical treatments and rare provider types	Performance where an error hurts the member most	$\geq 76\%$	140
Degraded wording	Robustness to imperfect entries	$\geq 85\%$	80
Out-of-scope files	Expected refusal	$\geq 99\%$	40

Thresholds are written before the results are known. A threshold set afterwards always adjusts itself to the figure obtained — and then it measures nothing at all.

The limits of this protocol

The 640 files come from the insurer's past. **A treatment that has never come up is not tested.** The evaluation measures performance against a history; it is not a guarantee about situations still to come, and the report repeats this.

Protocol v3 settled on ____ by ____ · next revision ____

Scope of this document: the service operated by Blue Lemon Agent. For an agent deployed at your premises, the applicable location is that of the architecture set out in the quotation and verified before commissioning; local execution is announced only for the configuration explicitly described and accepted in the quotation; the applicable isolation depends on the deployment mode set out in the quotation, no dedicated isolation being presumed; the encryption mechanisms in transit and at rest, their components and key management are those documented for the architecture chosen; roles and permissions are configured and accepted for the identities and systems actually connected; the events logged, their content, their retention period and who may access them are defined for the deployment chosen.