MELNYK, Oksana, POPOVYCH, Andriy, PURIJ, Nataliya, KOVALCHUK, Halyna, KOPKO, Iryna, LUPAK, Oksana, FIL, Vitalii, STYRANKA, Maksym and ZUKOW, Walery. Personification of the attributive effects of balneotherapy at Truskavets' spa. Pedagogy and Psychology of Sport. 2025;26:66805. eISSN 2450-6605.
https://doi.org/10.12775/PPS.2025.26.66805
https://apcz.umk.pl/PPS/article/view/66805

The journal has received 5 points in the Ministry of Science and Higher Education parametric evaluation. § 8. 2) and § 12. 1. 2) 22.02.2019.

The authors declare that there is no conflict of interest regarding the publication of this paper.
Received: 20.09.2025. Revised: 29.09.2025. Accepted: 07.10.2025. Published: 20.10.2025.

Personification of the attributive effects of balneotherapy at Truskavets' spa

Oksana I. Melnyk[1], Andriy I. Popovych[2], Nataliya A. Purij[3], Halyna Y. Kovalchuk[4], Iryna Y. Kopko[4], Oksana M. Lupak[4], Vitalii M. Fil[4], Maksym O. Styranka[5], Walery Zukow[6]

[1]Stepan Gzhytskyi National University of Veterinary Medicine and Biotechnologies, Lviv, UKRAINE omelnyk7@gmail.com
[2]Scientific group of Balneology of Hotel&Spa Complex "Karpaty", Truskavets', UKRAINE andriypopovych711@gmail.com
[3]Medical college, Boryslav, UKRAINE KosNatalya@ukr.net
[4]Ivan Franko Drohobych State Pedagogical University, Drohobych, UKRAINE
h.kovalchuk@dspu.edu.ua; kopkoiryna@dspu.edu.ua; oksana_lupak@dspu.edu.ua; fillvitalij@dspu.edu.ua
[5]IY Horbachevskyi National Medical University, Ternopil', UKRAINE
samsj3121017@gmail.com
[6]Nicolaus Copernicus University, Torun, POLAND w.zukow@wp.pl

OM: https://orcid.org/0000-0001-7928-4760
HK: https://orcid.org/0000-0002-5261-8422
IK: https://orcid.org/0000-0002-8607-7900
OL: https://orcid.org/0000-0002-1969-8643
VF: https://orcid.org/0000-0003-4770-5633
MS: https://orcid.org/0009-0003-1242-6738
WZ: https://orcid.org/0000-0002-7675-6117

Abstract
Background. It is generally accepted that the attributive effects of balneotherapy at the resort of Truskavets' are diuretic, cholecystokinetic, urolitholytic, and bactericidal. However, these effects do not occur in all patients, that is, they manifest themselves only as a tendency. Standing on the positions of personification and predictability, we set ourselves the goal: first, to identify a number of options for the attributive effects of balneotherapy at the resort of Truskavets', and secondly, to clarify the possibility of predicting one or another option. This article concerns the first stage of achieving the goal.
Materials and methods. Under an observations were 34 men and 10 women by age 24-76 years with chronic pyelonephritis and cholecystitis in the phase of remission. Testing was performed twice - on admission and after 7-10 days of standard balneotherapy (drinking of Naftussya bioactive water, applications of ozokerite, mineral pools). Daily diuresis and urine concentration of some components were determined with calculation of its lithogenicity, as well as indicators of cholecystokinetic activity and immunity with calculation of the bactericidal ability of blood neutrophils.
Results. In 10 patients (22.7%), a pronounced diuretic effect (D2+) was accompanied by a moderate cholecystokinetic effect (C+) and a moderate increase in bactericidal activity against E. coli (E+), but not Staph. aureus (A0) with a complete absence of changes in urine lithogenicity (L0). In 9 patients (20.5%), a moderate diuretic effect (D+) was accompanied by a pronounced cholecystokinetic effect (C2+) in combination with an increase in bactericidal activity against both E. coli (E+) and, to an even greater extent, Staph. aureus (A2+), while urine lithogenicity moderately increased (L+). A characteristic feature of the third cluster (8 patients, 18.2%) is a drastic cholecystokinetic effect (C3+) in combination with a moderate diuretic effect (D+) and a slight increase in

urine lithogenicity (L+) and the absence of significant changes in bactericidal activity against both types of microbes (E0&A0). The fourth cluster has only 3 patients in whom balneotherapy drastically suppressed cholekinetics (C3-), but significantly increased bactericidal activity against both types of microbes (E2+&A2+) with uncertain changes in diuresis and a slight increase in urine lithogenicity (L+). The last cluster was the most numerous (14 patients, 31.8%), characterized by slight anticholecystokinetic and antilithogenic effects and a moderate increase in bactericidal activity against Staph. aureus (A+) in the absence of changes in diuresis and bactericidal activity against E. coli.
Conclusion. Overall, the attributive effects of balneotherapy can be interpreted as physiologically favorable or at least neutral.
Keywords: Truskavets' spa, diuresis, urine lithogenicity, cholecystokinetics, immunity, personification, cluster and discriminant analyses.

INTRODUCTION

Historically, the first theory of the mechanism of therapeutic action of Naftussya bioactive water was diuretic-excretory [2,4,35,36]. The next attribute of its therapeutic action was considered to be the stimulation of choleresis and cholecystokinetic [5,34,35]. Attention was occasionally focused on the ability of Naftussya bioactive water to reduce the lithogenicity of both urine and bile [6,7,34]. The latest in this series, but the most important effect was the discovery of an increase in the bactericidal ability of the blood, which made it possible to explain the mechanism of suppression or even elimination of the inflammatory process in the urinary and hepatobiliary systems [10,28], which was previously explained by the washing out of microbes and pus by the accelerated flow of urine and bile. The apotheosis of the evolution of views on the mechanism of therapeutic action of Naftussya bioactive water was the theory of Popovych IL [26,27,29], xenobiotic-synbiotic according to the active factors and neuro-endocrine-immune adaptogenic according to the mechanism of their influence on the body. According to this theory, organic substances (xenobiotics) [11,12,15] and autochthonous bacteria (synbiotics) [14,34,35] have a modulating effect on the neuro-endocrine-immune complex, which, in turn, regulates both the activity of the urinary [6,7] and hepatobiliary [9,22,23] systems (as well as the cardiovascular system [30]), and the processes of inflammation, hematopoiesis, hemostasis [19,20,32], carcinogenesis [13], etc.

It should be noted that in the aforementioned works the authors frankly cited the facts that diuretic, cholecystokinetic, litholytic, and bactericidal effects do not occur in all patients, that is, they manifest themselves only as a tendency [3]. Standing on the positions of personification and predictability, we set ourselves the goal: first, to identify a number of options for the attributive diuretic, cholecystokinetic, urolitholytic, and bactericidal effects of balneotherapy at the resort of Truskavets', and secondly, to clarify the possibility of predicting one or another option.

This article concerns the first stage of achieving the goal.

RESEARCH OBJECTIVE

The primary objective of this investigation is to identify distinct variants of attributive effects of balneotherapy at Truskavets' spa, specifically examining diuretic, cholecystokinetic, urolitholytic, and bactericidal effects, and to determine the possibility of predicting specific variants of therapeutic response in patients with chronic pyelonephritis and cholecystitis in the phase of remission. This study represents the first stage of a comprehensive research program aimed at personification and predictability of balneotherapeutic interventions, moving beyond the traditional approach of describing effects as mere tendencies toward a more individualized understanding of therapeutic responses. The investigation seeks to elucidate whether the historically recognized attributive effects of Naftussya bioactive water and associated balneotherapeutic modalities manifest uniformly across all patients or whether distinct response patterns can be identified, characterized, and ultimately predicted based on baseline clinical, biochemical, and immunological parameters.

RESEARCH PROBLEMS

The first research problem addresses the fundamental question of therapeutic response heterogeneity: do the attributive effects of balneotherapy, including diuretic, cholecystokinetic, urolitholytic, and bactericidal effects, manifest uniformly in all patients, or are there qualitatively and quantitatively different variants of therapeutic response that can be systematically identified and characterized? This problem stems from clinical observations suggesting that while balneotherapy at Truskavets' spa demonstrates consistent trends at the population level, individual patient responses vary considerably, with some patients exhibiting pronounced diuretic effects while others show minimal changes in urine output but substantial improvements in gallbladder motility or bactericidal capacity. Understanding this heterogeneity is crucial for moving from population-based to personalized balneotherapy protocols.

The second research problem concerns the personification of balneotherapy effects: what are the characteristic profiles of clinical parameter changes in different patient groups, and can specific clusters of response to balneotherapy be identified based on multivariate analysis of diuresis, cholecystokinetic index, urine lithogenicity, and bactericidal capacity against Gram-positive and Gram-negative bacteria? This problem requires sophisticated statistical approaches, including cluster analysis and discriminant function analysis, to identify natural groupings of patients based on their response patterns. The identification of such clusters would represent a significant advancement in balneology, providing a framework for understanding the mechanisms underlying different response patterns and potentially guiding treatment individualization.

The third research problem examines the dependence of therapeutic effects on baseline clinical state: to what extent do initial values of clinical parameters, including diuresis, cholecystokinetic index, urine lithogenicity, and bactericidal activity against *Escherichia coli* and *Staphylococcus aureus*, determine the direction and intensity of changes after balneotherapy? This problem is particularly relevant from a clinical perspective, as it addresses whether balneotherapy exerts primarily normalizing effects, correcting pathological deviations while leaving normal parameters unchanged, or whether it produces systematic shifts regardless of baseline values. Understanding these relationships would inform patient selection criteria and help establish realistic therapeutic expectations based on pre-treatment assessment.

The fourth research problem investigates interactions between therapeutic effects: are there specific patterns of co-occurrence or mutual exclusion of individual balneotherapy effects, such that a strong diuretic effect correlates with or precludes certain profiles of cholecystokinetic, urolitholytic, or bactericidal changes? This problem addresses the potential compensatory or synergistic relationships between different physiological systems responding to balneotherapy. For instance, it remains unclear whether the neuro-endocrine-immune mechanisms underlying enhanced diuresis operate independently of or in coordination with those regulating gallbladder motility and bile secretion. Elucidating these interactions would provide insights into the integrative physiological responses to balneotherapy and might reveal trade-offs or complementarities in therapeutic effects.

The fifth research problem concerns the prediction of therapeutic response: is it possible to predict a patient's membership in a specific response cluster based on baseline parameters, including biochemical markers, immunological indicators, and functional assessments, and if so, which parameters possess the greatest discriminatory power? This problem represents the ultimate goal of personalized balneotherapy, as accurate prediction of response patterns would enable clinicians to tailor treatment protocols, set appropriate expectations, and potentially modify interventions to optimize outcomes for individual patients. The development of a robust predictive model would require identification of the minimal set of baseline parameters necessary for accurate classification while maintaining clinical feasibility and cost-effectiveness.

RESEARCH HYPOTHESES

The first research hypothesis posits that the effects of balneotherapy at Truskavets' spa manifest as at least three to five distinct response clusters, each characterized by specific profiles of changes in diuresis, cholecystokinetics, urine lithogenicity, and bactericidal activity parameters, with no single cluster comprising more than 35% of the studied population. This hypothesis is grounded in the clinical observation that patients with similar diagnoses and disease stages nevertheless exhibit considerable variability in their responses to balneotherapy, suggesting the existence of distinct response phenotypes. The hypothesis predicts that cluster analysis will reveal natural groupings of patients based on their multivariate response patterns, with each cluster representing a qualitatively different mode of physiological adaptation to balneotherapeutic interventions. The specification that no cluster should exceed 35% of the population reflects the expectation of genuine heterogeneity rather than a dominant response pattern with minor variations. This hypothesis challenges the traditional view of balneotherapy effects as uniform tendencies and proposes instead a model of discrete response variants, each potentially reflecting different underlying mechanisms of action or different states of neuro-endocrine-immune regulation.

The second research hypothesis proposes an inverse relationship between the intensity of diuretic and cholecystokinetic effects, specifically that patients with pronounced increase in diuresis exceeding 20% of baseline values will show weaker changes in cholecystokinetics, defined as less than 15% change in cholecystokinetic index, and vice versa. This hypothesis is based on physiological considerations regarding the finite capacity of neuro-endocrine-immune regulatory systems and the potential for competitive allocation of adaptive resources between different organ systems. The hypothesis suggests that the mechanisms underlying enhanced renal function and those regulating hepatobiliary function may operate in a compensatory rather than synergistic manner, possibly due to shared regulatory pathways or limited availability of key mediators. If confirmed, this hypothesis would have important implications for treatment planning, suggesting that maximizing one therapeutic effect might come at the expense of another, and that balanced rather than maximal stimulation might be preferable in patients requiring improvements in multiple systems. The hypothesis also implies that the traditional concept of

comprehensive balneotherapy effects occurring simultaneously in all target systems may need revision in favor of a more nuanced understanding of selective or prioritized responses.

The third research hypothesis asserts that balneotherapy exerts a normalizing effect, whereby parameters deviating from normal ranges before treatment undergo statistically significant correction toward reference values with $p<0.05$, while parameters within normal range remain stable with changes less than 10% of baseline values. This hypothesis reflects the concept of balneotherapy as an adaptogenic intervention that enhances homeostatic regulation rather than producing uniform directional changes regardless of initial state. The hypothesis predicts that patients with elevated urine lithogenicity will experience reductions, those with suppressed bactericidal activity will show increases, and those with gallbladder dyskinesia will move toward normal motility patterns, while patients with normal baseline values will not experience significant shifts. This normalizing effect would be consistent with the neuro-endocrine-immune adaptogenic theory of balneotherapy action proposed by Popovych and colleagues, which emphasizes the role of bioactive water components in modulating regulatory systems rather than directly affecting target organs. Confirmation of this hypothesis would support the use of balneotherapy as a safe intervention with built-in physiological limits, reducing concerns about over-stimulation or excessive correction of physiological parameters.

The fourth research hypothesis proposes that the increase in bactericidal activity of neutrophils has a selective character with respect to microorganism type, specifically that in patients with reduced baseline activity against *Escherichia coli*, bactericidal capacity will increase by at least 25%, while activity against *Staphylococcus aureus* will remain unchanged or increase to a lesser extent, defined as less than 15%. This hypothesis is based on immunological considerations regarding the different mechanisms of neutrophil-mediated killing of Gram-negative versus Gram-positive bacteria and the possibility that balneotherapy preferentially enhances certain antimicrobial pathways. *Escherichia coli*, as a Gram-negative organism, is particularly relevant to urinary tract infections common in pyelonephritis, while *Staphylococcus aureus*, a Gram-positive organism, represents a different immunological challenge. The hypothesis suggests that the immunomodulatory effects of balneotherapy may be targeted toward the most clinically relevant pathogens in the context of urological disorders, or alternatively, that the mechanisms enhanced by balneotherapy are more effective against Gram-negative bacteria. This selective enhancement would represent a sophisticated adaptive response rather than non-specific immune stimulation, and if confirmed, would provide insights into the specific immunological pathways activated by balneotherapy.

The fifth research hypothesis states that a discriminant model based on at least 15 baseline parameters, including biochemical, immunological, and functional indicators, will allow retrospective classification of patients into appropriate response clusters with accuracy of at least 85%, with cholecystokinetic parameters showing the greatest discriminatory power as evidenced by Wilks' Lambda less than 0.10 and $p<0.001$. This hypothesis represents the culmination of the personification approach, proposing that pre-treatment assessment can reliably predict post-treatment response patterns. The specification of 15 parameters reflects a balance between comprehensive assessment and clinical practicality, while the 85% accuracy threshold represents a clinically meaningful level of predictive power that would justify the implementation of personalized treatment protocols. The hypothesis that cholecystokinetic parameters will show the greatest discriminatory power is based on preliminary observations suggesting that baseline gallbladder function may be particularly informative regarding overall neuro-endocrine-immune regulatory capacity. If confirmed, this hypothesis would enable development of a practical clinical tool for treatment individualization, allowing clinicians to predict which patients will respond with primarily diuretic effects, which will show cholecystokinetic responses, and which will exhibit enhanced bactericidal activity, thereby enabling targeted treatment planning and appropriate patient counseling.

STATISTICAL HYPOTHESES

Statistical Hypothesis 1: Significance of Inter-cluster Differences. The null hypothesis states that there are no statistically significant differences in mean values of changes in attributive parameters, specifically diuresis, cholecystokinetic index, urine lithogenicity, bactericidal capacity of neutrophils against *Escherichia coli*, and bactericidal capacity against *Staphylococcus aureus*, between identified patient clusters. The alternative hypothesis proposes that there are statistically significant differences in mean values of changes in at least three of five attributive parameters between clusters, as assessed by analysis of variance or multivariate analysis of variance with $p<0.05$ after Bonferroni correction for multiple comparisons. The operationalization of this hypothesis involves multivariate analysis of variance with cluster membership as the independent variable with five levels corresponding to the identified clusters, and the five change scores in attributive parameters as dependent variables. The criterion for rejection of the null hypothesis is set at Wilks' Lambda less than 0.10, F-statistic greater than 10, and $p<0.001$, representing a very stringent threshold that accounts for the exploratory nature of cluster identification. Post-hoc comparisons between specific cluster pairs will be conducted using Tukey's Honestly

Significant Difference test to control family-wise error rate while identifying which specific clusters differ significantly from one another. The statistical power of this test is enhanced by the multivariate approach, which considers the correlational structure among dependent variables and is more sensitive to group differences than multiple univariate tests. This hypothesis is fundamental to validating the cluster solution, as significant inter-cluster differences in response parameters are necessary to justify the interpretation of clusters as representing distinct response phenotypes rather than arbitrary divisions of a continuous distribution. The effect sizes will be reported using partial eta-squared for individual parameters and Pillai's trace for the overall multivariate effect, providing information about the practical significance of differences in addition to statistical significance. Assumptions of MANOVA, including multivariate normality, homogeneity of variance-covariance matrices, and absence of multicollinearity, will be assessed using Mardia's test, Box's M test, and variance inflation factors respectively, with appropriate transformations or robust methods employed if assumptions are violated.

Statistical Hypothesis 2: Discriminatory Power of the Model. The null hypothesis states that the discriminant analysis model based on baseline parameters does not allow for significantly better than random classification of patients into response clusters, with accuracy not exceeding 25%, which represents chance-level performance for classification into five groups. The alternative hypothesis proposes that the discriminant analysis model achieves classification accuracy of at least 80%, significantly higher than random as evidenced by chi-square statistic greater than 50 with $p<0.001$, and that the first canonical discriminant root explains at least 85% of between-group variance. The operationalization involves stepwise discriminant analysis with 20 to 30 baseline parameters as potential predictor variables, including biochemical markers such as serum and urinary concentrations of electrolytes, metabolites, and lithogenic components; immunological parameters including lymphocyte subpopulations, immunoglobulin levels, and phagocytic function indices; and functional assessments including baseline diuresis and cholecystokinetic index. The stepwise selection procedure will use F-to-enter greater than 2.0 with $p<0.10$ as the inclusion criterion and F-to-remove less than 1.5 with $p>0.15$ as the exclusion criterion, balancing the goals of model parsimony and predictive accuracy. Model validation will be assessed through the classification matrix showing the number and percentage of correctly classified cases for each cluster, with overall accuracy, sensitivity, and specificity calculated for each cluster. Cohen's kappa coefficient will be computed to assess agreement between predicted and actual cluster membership while accounting for chance agreement, with values greater than 0.75 considered excellent agreement. The discriminant functions will be evaluated using canonical correlation coefficients, with values greater than 0.90 indicating very strong relationships between predictor variables and cluster membership, and eigenvalues greater than 10 indicating substantial separation between clusters in the discriminant space. The relative importance of individual predictor variables will be assessed using standardized discriminant function coefficients and structure coefficients, which indicate the correlation between each predictor and the discriminant functions. Cross-validation will be performed using leave-one-out classification to assess the stability and generalizability of the model, with the expectation that cross-validated accuracy should not fall below 75% if the model is robust. This hypothesis is critical for establishing the clinical utility of personalized balneotherapy, as accurate prediction of response patterns based on readily obtainable baseline parameters would enable truly individualized treatment planning.

Statistical Hypothesis 3: Parameter Correlations Within Clusters. The null hypothesis states that there are no significant correlations, defined as absolute values of correlation coefficients less than 0.30 with $p>0.05$, between changes in individual attributive parameters within identified clusters. The alternative hypothesis proposes that in at least three clusters, there are strong correlations with absolute values of correlation coefficients of at least 0.60 and $p<0.01$ between changes in at least two pairs of attributive parameters, and furthermore, that the direction of correlation differs between clusters, indicating cluster-specific patterns of parameter relationships. The operationalization involves computation of correlation matrices separately for each of the five clusters, with Pearson correlation coefficients calculated for pairs of variables that meet assumptions of bivariate normality and linearity, and Spearman rank correlation coefficients calculated for pairs that violate these assumptions as assessed by Shapiro-Wilk tests and scatter plot inspection. Given the multiple comparisons involved in examining all possible pairs of five parameters across five clusters, correction for false discovery rate will be applied using the Benjamini-Hochberg procedure, which controls the expected proportion of false positives among rejected null hypotheses while maintaining greater statistical power than family-wise error rate control methods such as Bonferroni correction. The analysis will specifically examine correlations between diuresis changes and cholecystokinetic changes to test the compensatory hypothesis, between urine lithogenicity changes and diuresis changes to assess the dilution effect, between bactericidal capacity changes against the two bacterial species to evaluate specificity of immune enhancement, and between cholecystokinetic changes and bactericidal capacity changes to explore potential links between hepatobiliary function and immune status. Heat maps with hierarchical clustering dendrograms will be constructed to visualize the correlation structures within each cluster, facilitating identification of distinct patterns of parameter relationships. The hypothesis that correlation directions differ

between clusters will be tested using Fisher's r-to-z transformation to compare correlation coefficients between clusters, with significant differences indicating that the relationships between parameters are moderated by cluster membership. This hypothesis addresses the question of whether the attributive effects of balneotherapy operate independently or in coordinated fashion, and whether these relationships differ between response phenotypes, providing insights into the integrative physiological mechanisms underlying different patterns of therapeutic response.

Statistical Hypothesis 4: Dependence of Effects on Baseline Values. The null hypothesis states that there is no statistically significant relationship between baseline parameter values and the magnitude of their changes after balneotherapy, as evidenced by coefficient of determination less than 0.10 and $p>0.05$ in regression analysis. The alternative hypothesis proposes that baseline parameter values explain at least 30% of variance in their changes after balneotherapy, indicated by R-squared of at least 0.30 with $p<0.01$, and furthermore, that the relationship has a non-linear character, with quadratic or logarithmic regression models providing significantly better fit than linear models as evidenced by change in R-squared of at least 0.10 with $p<0.05$ for the additional terms. The operationalization involves multiple regression analysis with five separate models, one for each attributive parameter change as the dependent variable and the corresponding baseline value as the primary independent variable, along with other potentially relevant baseline parameters as covariates. For each dependent variable, three models will be fitted: a simple linear model with only the baseline value of that parameter as predictor, a quadratic model adding the squared term of the baseline value, and a logarithmic model using the natural logarithm of the baseline value. Model comparison will be conducted using F-tests for nested models to assess whether the additional terms in non-linear models significantly improve prediction, and information criteria including Akaike Information Criterion and Bayesian Information Criterion will be calculated to compare models while penalizing complexity. Diagnostic procedures will include examination of residual plots to assess homoscedasticity and normality of errors, calculation of variance inflation factors to detect multicollinearity among predictors, identification of influential observations using Cook's distance and DFBETAS, and assessment of model assumptions using appropriate statistical tests. The hypothesis that non-linear models provide better fit is based on physiological reasoning suggesting that the potential for change is greatest when baseline values are most deviant from normal, with diminishing returns as values approach optimal levels, creating a curvilinear relationship. Separate analyses will be conducted for patients with baseline values above versus below the reference range to test whether balneotherapy effects are directionally appropriate, that is, whether elevated parameters decrease and suppressed parameters increase. This hypothesis directly addresses the concept of balneotherapy as a normalizing intervention and has important implications for patient selection and outcome prediction, as it would indicate that the magnitude of therapeutic benefit depends critically on the degree of baseline dysfunction.

Statistical Hypothesis 5: Stability of Cluster Structure. The null hypothesis states that the cluster structure is not stable, with changes in clustering method or number of clusters leading to reclassification of more than 40% of patients, corresponding to kappa agreement coefficient less than 0.40, indicating poor agreement. The alternative hypothesis proposes that the five-cluster structure is stable, with alternative clustering methods including hierarchical clustering with Ward's linkage and complete linkage, partitioning around medoids, and density-based spatial clustering of applications with noise, as well as solutions with four to six clusters, showing high agreement with the basic k-means five-cluster solution as evidenced by kappa coefficient greater than 0.70, Adjusted Rand Index greater than 0.75, and Normalized Mutual Information greater than 0.70. The operationalization begins with the basic k-means clustering solution with $k=5$, which serves as the reference classification. Hierarchical clustering will be performed using Ward's minimum variance method, which minimizes within-cluster sum of squares, and complete linkage method, which minimizes maximum distance between cluster members, with dendrograms inspected to determine the optimal number of clusters based on the pattern of fusion coefficients and visual assessment of cluster separation. Partitioning around medoids, a more robust alternative to k-means that uses actual observations as cluster centers rather than computed means, will be applied to assess whether the cluster structure depends on the specific algorithmic approach. Silhouette analysis will be conducted for solutions with three to seven clusters, computing the average silhouette width for each solution, with higher values indicating better-defined clusters; this analysis will determine whether five clusters is indeed optimal or whether a different number of clusters provides better separation. For each alternative clustering solution, agreement with the reference k-means five-cluster solution will be quantified using Cohen's kappa coefficient, which corrects for chance agreement; Adjusted Rand Index, which measures similarity between two clusterings while adjusting for chance grouping; and Normalized Mutual Information, which quantifies the mutual dependence between cluster assignments. The criterion for stability is that all three agreement measures exceed 0.70 when comparing alternative methods with five clusters to the reference solution, and that the five-cluster solution shows higher agreement with the reference than do four-cluster or six-cluster solutions. Additionally, stability will be assessed through bootstrap resampling, generating 1000 bootstrap samples and performing k-means clustering on each, then

computing the adjusted Rand Index between each bootstrap solution and the original solution, with mean ARI greater than 0.75 indicating stable cluster structure. This comprehensive assessment of cluster stability is essential for establishing confidence in the identified response phenotypes, as unstable cluster solutions would suggest that the observed groupings are artifacts of the particular clustering method or sample rather than representing genuine distinct response patterns. Stable cluster structure would support the validity of the personification approach and justify the development of clinical tools based on cluster membership.

Statistical Hypothesis 1: Significance of Inter-cluster Differences
H_0: There are no statistically significant differences in mean values of changes in attributive parameters (diuresis, cholecystokinetic index, urine lithogenicity, BCCN vs *E. coli*, BCCN vs *Staph. aureus*) between identified patient clusters.
H_1: There are statistically significant differences in mean values of changes in at least 3 of 5 attributive parameters between clusters (ANOVA/MANOVA, $p<0.05$ with Bonferroni correction).
Operationalization:
Test: Multivariate Analysis of Variance (MANOVA)
Dependent variables: Δdiuresis, Δcholecystokinetic index, Δlithogenicity, ΔBCCN *E.coli*, ΔBCCN *S.aureus*
Independent variable: cluster membership (5 levels)
Criterion: Wilks' Lambda <0.10, F>10, $p<0.001$
Post-hoc: Tukey's HSD tests for pairwise comparisons

Statistical Hypothesis 2: Discriminatory Power of the Model
H_0: The discriminant analysis model based on baseline parameters does not allow for significantly better than random classification of patients into response clusters (accuracy ≤25%, i.e., at chance level for 5 clusters).
H_1: The discriminant analysis model achieves classification accuracy ≥80% (significantly higher than random, $\chi^2>50$, $p<0.001$), with the first canonical root explaining ≥85% of between-group variance.
Operationalization:
Test: Stepwise discriminant analysis
Predictor variables: 20-30 baseline parameters
Inclusion criteria: F to enter >2.0, $p<0.10$
Validation: classification matrix, Cohen's kappa coefficient >0.75
Evaluation: Canonical R >0.90, Eigenvalue >10

Statistical Hypothesis 3: Parameter Correlations Within Clusters
H_0: There are no significant correlations ($|r|<0.30$, $p>0.05$) between changes in individual attributive parameters within identified clusters.
H_1: In at least 3 clusters, there are strong correlations ($|r|\geq0.60$, $p<0.01$) between changes in at least two pairs of attributive parameters, with the direction of correlation differing between clusters.
Operationalization:
Test: Pearson correlation (for normal distributions) or Spearman (for non-parametric)
Analysis: correlation matrices for each cluster separately
Correction: False Discovery Rate (FDR) for multiple comparisons
Visualization: correlation heat maps with dendrograms

Statistical Hypothesis 4: Dependence of Effects on Baseline Values
H_0: There is no statistically significant relationship between baseline parameter values and the magnitude of their changes after balneotherapy ($R^2<0.10$, $p>0.05$ in regression analysis).
H_1: Baseline parameter values explain ≥30% of variance in their changes after balneotherapy ($R^2\geq0.30$, $p<0.01$), with the relationship being non-linear (quadratic or logarithmic model significantly better than linear, $\Delta R^2\geq0.10$, $p<0.05$).
Operationalization:
Test: multiple regression analysis (linear, quadratic, logarithmic)
Dependent variables: Δ parameters (5 models)
Independent variables: baseline parameter values
Model comparison: F-test for R^2 difference, AIC, BIC
Diagnostics: residual analysis, VIF for multicollinearity

Statistical Hypothesis 5: Stability of Cluster Structure
H_0: The cluster structure is not stable - changing the clustering method or number of clusters leads to reclassification of >40% of patients (kappa agreement coefficient <0.40).
H_1: The 5-cluster structure is stable - alternative clustering methods (hierarchical, PAM, DBSCAN) and solutions with 4-6 clusters show high agreement with the basic solution (kappa >0.70, Adjusted Rand Index >0.75).
Operationalization:
Basic method: k-means clustering (k=5)
Validation methods:
Hierarchical clustering (Ward, complete linkage)
Partitioning Around Medoids (PAM)
Silhouette analysis for k=3-7
Agreement measures:
Cohen's kappa coefficient
Adjusted Rand Index (ARI)
Normalized Mutual Information (NMI)
Criterion: all measures >0.70 for k=5

SUMMARY OF STATISTICAL METHODOLOGY
Statistical Methods Applied in the Study:
Cluster Analysis (k-means) - identification of 5 clusters
Discriminant Analysis (stepwise) - predictive model
MANOVA - between-group comparisons
Correlation Analysis - relationships between parameters
Z-score Normalization - comparability of effects
Key Statistical Results from the Document:
Wilks' Lambda: 0.00032 ($p<10^{-6}$)
Classification accuracy: 100% (retrospective)
Canonical R: 0.994 (Root 1), 0.934 (Root 2)
F-value: 13.0 ($p<10^{-6}$)
Identified Clusters:
Cluster A (22.7%): D2+C+L0E+A0 - Strong diuretic, moderate cholecystokinetic
Cluster B (20.5%): D+C2+L+E+A2+ - Moderate diuretic, strong cholecystokinetic
Cluster C (18.2%): D+C3+L+E0A0 - Drastic cholecystokinetic effect
Cluster D (6.8%): D0C3-L+E2+A2+ - Anticholecystokinetic, strong bactericidal
Cluster E (31.8%): D0C-L-E0A+ - Slight anticholecystokinetic, antilithogenic

DISCUSSION OF STATISTICAL METHODOLOGY
The statistical approach employed in this investigation represents a sophisticated integration of exploratory and confirmatory multivariate techniques designed to identify, characterize, and validate distinct patterns of response to balneotherapy while establishing the feasibility of predicting response patterns from baseline assessments. The initial application of k-means cluster analysis to normalized effect sizes addresses the fundamental question of whether patients can be meaningfully grouped based on their multivariate response profiles, with the choice of five clusters representing a balance between parsimony and adequate representation of response heterogeneity. The normalization of effect sizes using Z-score transformation is critical for ensuring that parameters measured on different scales contribute appropriately to cluster formation, preventing variables with larger absolute values from dominating the clustering algorithm. The subsequent application of discriminant function analysis serves dual purposes: first, to identify which baseline parameters most strongly differentiate between response clusters, thereby providing insights into the mechanisms underlying different response phenotypes; and second, to develop a predictive model that could be applied prospectively to classify new patients into expected response categories before treatment initiation. The stepwise selection procedure in discriminant analysis balances the competing goals of including all relevant predictors while avoiding overfitting and maintaining model parsimony, with the relatively liberal inclusion criterion of $p<0.10$ justified by the exploratory nature of predictor identification in this initial investigation. The achievement of 100% retrospective classification accuracy, while impressive, must be interpreted cautiously as it may reflect overfitting, particularly given the relatively small sample size of 44 patients distributed across five clusters; this concern motivates the emphasis on cross-validation and the proposed stability analyses in Statistical Hypothesis 5. The multivariate analysis of variance approach in Statistical Hypothesis 1 provides a rigorous test of whether the identified clusters truly differ in their response profiles, with the multivariate

test being more appropriate than multiple univariate tests because it accounts for correlations among the dependent variables and controls Type I error rate. The correlation analyses proposed in Statistical Hypothesis 3 address the important question of whether different attributive effects operate independently or in coordinated fashion, with the cluster-specific approach allowing for the possibility that these relationships differ between response phenotypes, potentially reflecting different underlying mechanisms or regulatory states. The regression analyses in Statistical Hypothesis 4 directly test the normalizing effect hypothesis by examining whether the magnitude and direction of changes depend on baseline values, with the inclusion of non-linear models reflecting the physiologically plausible expectation that the relationship between baseline state and response magnitude may not be linear. The comprehensive stability assessment in Statistical Hypothesis 5, incorporating multiple clustering algorithms, multiple cluster numbers, multiple agreement indices, and bootstrap resampling, addresses the critical question of whether the identified cluster structure represents genuine distinct response phenotypes or merely reflects arbitrary divisions of continuous variation. The convergence of evidence from multiple statistical approaches, each addressing different aspects of the personification hypothesis, provides a robust framework for establishing the validity and clinical utility of individualized balneotherapy based on predicted response patterns. The statistical power of the analyses is enhanced by the multivariate approach, which leverages the correlational structure among variables to detect effects that might not be apparent in univariate analyses, although the relatively modest sample size of 44 patients necessitates cautious interpretation and emphasizes the need for replication in larger samples. The integration of effect size measures alongside significance tests ensures that statistical significance is not confused with clinical or practical significance, with effect sizes providing information about the magnitude of differences and relationships that is essential for clinical decision-making. The proposed statistical framework thus provides a comprehensive approach to testing the personification hypothesis while addressing potential limitations and alternative explanations, establishing a rigorous foundation for advancing balneotherapy from a one-size-fits-all approach toward truly personalized medicine based on predicted individual response patterns.

MATERIALS AND METHODS

Under an observations were 34 men and 10 women by age 24-76 years with chronic pyelonephritis and cholecystitis in the phase of remission. Testing was performed twice - on admission and after 7-10 days of standard balneotherapy (drinking of Naftussya bioactive water, applications of ozokerite, mineral pools) [24,25,28].

First, daily urine was collected, in which was determined the concentration of electrolytes: calcium (by reaction with arsenase III), magnesium (by reaction with colgamite), phosphates (phosphate-molybdate method), chloride (mercury-rhodanidine method), sodium and potassium (flamming photometry) as well as nitric metabolites: creatinine (by Jaffe's color reaction by Popper's method), urea (urease method by reaction with phenolhypochlorite), and uric acid (uricase method). Urine lithogenicity index (Lith) was also calculated by the Tiselius' HS [33] formula modifed by Flyunt VR et al [6,7]:

$$Lith = (Uric\ acid \bullet Calcium/Magnesium \bullet Creatinine)^{0,25}.$$

The same metabolic parameters were determined in serum. The analysis carried out according to instructions [8] with the use of analyzers "Reflotron" (BRD) and "Pointe-180" (USA) and corresponding sets of reagents.

On the tone and motility of gall-bladder judged by its volume on an empty stomach in the morning and after 5, 15 and 30 min after ingestion of cholekinetic (50 ml of 40% solution of xylitol). The method echoscopy (echocamera "Radmir") applicated. To quantify cholekinetics index, the area between the cholecystovolumogram and the basal line was calculated [22,23].

Parameters of phagocytic function of neutrophils estimated as described by Kovbasnyuk MM [20,29]. The objects of phagocytosis served daily cultures of Staphylococcus aureus (ATCC N 25423 F49) as typical specimen for Gram-positive Bacteria, and Escherichia coli (O55 K59) as typical representative of Gram-negative Bacteria. Take into account the following parameters of Phagocytosis: activity (percentage of neutrophils, in which found microbes - Hamburger's Phagocytic Index PhI), intensity (number of microbes absorbed one phagocytes - Microbial Count MC or Right's Index) and completeness (percentage of dead microbes - Killing Index KI). On the basis of the registered partial parameters of phagocytosis, taking into account the content of neutrophils (N) in 1 L of blood, the integral parameter - the bactericidal capacity of neutrophils - was calculated by the equation:

$$BCCN\ (10^9\ Bact/L) = N\ (10^9/L) \bullet PhI\ (\%) \bullet MC\ (Bact/Phag) \bullet KI\ (\%) \bullet 10^{-4}.$$

Parameters of cellular and humoral links of Immunity evaluated as described in the manuals [17,21]. For phenotyping subpopulations of lymphocytes used the methods of rosette formation with sheep erythrocytes on which adsorbed monoclonal antibodies against receptors CD3, CD4, CD8, CD22 and CD56 from company "Granum" (Kharkiv) with visualization under light microscope with immersion system. Subpopulation of T cells with receptors high affinity determined by test of "active" rosette formation. The state of humoral immunity judged

by the concentration in serum of Immunoglobulins of classes G, A, M (ELISA, analyser “Immunochem”, USA) and circulating immune complexes (by polyethylene glycol precipitation method).
Normal (reference) values of variables are taken from the instructions and/or database of the Truskavetsian Scientific School of Balneology [29].
For statistical analysis used the software package "Statistica 6.4".

Study Design
A prospective observational study with repeated measurements (before and after balneotherapeutic intervention) was conducted on a cohort of 44 patients (34 men and 10 women) aged 24-76 years with chronic pyelonephritis and cholecystitis in the remission phase. The study employed a pre-post design with assessments performed at baseline (upon admission to Truskavets' spa) and after 7-10 days of standard balneotherapy, which included drinking Naftussya bioactive water, ozokerite applications, and mineral pool bathing.

Data Collection and Variables
Primary outcome variables included five attributive effects of balneotherapy: (1) daily diuresis volume measured in milliliters per 24 hours; (2) cholecystokinetic index calculated from gallbladder volume measurements before and after choleretic stimulus; (3) urine lithogenicity index computed from concentrations of lithogenic and antilithogenic components; (4) bactericidal capacity of neutrophils against Escherichia coli measured by the reduction in viable bacterial count after incubation with patient blood; and (5) bactericidal capacity of neutrophils against Staphylococcus aureus using the same methodology. **Secondary variables** included approximately 20-30 baseline biochemical, immunological, and functional parameters that served as potential predictors of response patterns.

Statistical Analysis Methods
1. Data Normalization
All effect sizes were standardized using Z-score transformation to enable comparison of parameters measured on different scales and to prevent variables with larger absolute values from dominating multivariate analyses. The Z-score for each parameter change was calculated as $Z = (X - \mu) / \sigma$, where X represents the observed change value, μ is the mean change across all patients, and σ is the standard deviation of changes. This transformation converts all variables to a common scale with mean of zero and standard deviation of one, facilitating interpretation of effect magnitudes and ensuring equal weighting in clustering algorithms.
2. Cluster Analysis
K-means clustering algorithm was applied to the five normalized effect size variables to identify distinct patterns of response to balneotherapy. The number of clusters was set at k=5 based on theoretical considerations regarding the expected heterogeneity of responses and preliminary analysis of within-cluster sum of squares for solutions ranging from k=3 to k=8. The algorithm iteratively assigns patients to clusters and recalculates cluster centroids to minimize within-cluster variance while maximizing between-cluster separation. Convergence was achieved when cluster assignments stabilized across iterations. The resulting five clusters were characterized by their centroid values on each of the five attributive effect dimensions, and patients were assigned to the cluster with the nearest centroid based on Euclidean distance in the five-dimensional standardized space.
Cluster distribution: Cluster A comprised 10 patients (22.7%) characterized by strong diuretic effect with moderate cholecystokinetic effect; Cluster B included 9 patients (20.5%) with moderate diuretic and strong cholecystokinetic effects; Cluster C contained 8 patients (18.2%) with drastic cholecystokinetic effect; Cluster D consisted of 3 patients (6.8%) with anticholecystokinetic but strong bactericidal effects; and Cluster E was the largest with 14 patients (31.8%) showing subtle normalizing effects.
3. Discriminant Function Analysis
Stepwise discriminant analysis was performed to identify baseline parameters that optimally differentiate between the five response clusters and to develop a predictive classification model. The analysis proceeded in two stages: an initial model with four baseline variables (cholecystokinetic index, urine lithogenicity index, bactericidal capacity against both bacterial species) and an extended model incorporating approximately 20 baseline biochemical, immunological, and functional parameters.
Variable selection criteria: Forward stepwise selection was employed with F-to-enter threshold of 2.0 ($p<0.10$) for inclusion and F-to-remove threshold of 1.5 ($p>0.15$) for exclusion. The relatively liberal inclusion criterion was justified by the exploratory nature of predictor identification in this initial investigation.
Initial model results: The four-variable model achieved Wilks' Lambda of 0.0392, $F(16,1) = 13.0$, $p<10^{-6}$, with retrospective classification accuracy of 97.7%. Baseline cholecystokinetic index showed the strongest discriminatory power (Wilks' Lambda = 0.604, Partial Lambda = 0.065, F-remove = 129, $p<10^{-6}$), followed by

urine lithogenicity index, bactericidal capacity against Staphylococcus aureus, and bactericidal capacity against Escherichia coli.

Extended model results: The 20-variable model achieved Wilks' Lambda of 0.00032, $F(89) = 6.8$, $p<10^{-6}$, with perfect retrospective classification accuracy of 100%. The model included baseline biochemical parameters (urine concentrations of urea, chloride, calcium, phosphate, magnesium, creatinine; serum uric acid), immunological parameters (eosinophil percentage, serum IgG, phagocytic indices, lymphocyte subpopulations CD3+, CD4+, CD8+, CD22+), and functional parameters (cholecystokinetic index, urine lithogenicity index, fasting gallbladder volume, daily diuresis).

Canonical discriminant functions: Four canonical roots were extracted. Root 1 had eigenvalue of 81.2, explaining 89.6% of variance with canonical correlation of 0.994 (Wilks' Lambda = 0.00032, $\chi^2 = 289$, df = 89, $p<10^{-6}$). Root 2 had eigenvalue of 6.80, explaining 7.6% of variance with canonical correlation of 0.934 (Wilks' Lambda = 0.0089, $\chi^2 = 169$, df = 66, $p<10^{-6}$). Root 3 had eigenvalue of 1.90, explaining 2.1% of variance with canonical correlation of 0.810 (Wilks' Lambda = 0.0642, $\chi^2 = 96$, df = 45, $p<10^{-6}$). Root 4 had eigenvalue of 0.60, explaining 0.7% of variance with canonical correlation of 0.613 (Wilks' Lambda = 0.6250, $\chi^2 = 16$, df = 26, $p = 0.92$, not significant). The first three roots collectively explained 99.3% of between-group variance, indicating excellent discriminatory power.

4. Multivariate Analysis of Variance (MANOVA)

MANOVA was conducted to test the statistical significance of differences between the five clusters in their multivariate response profiles. Cluster membership served as the independent variable (five levels), while the five attributive effect changes (diuresis, cholecystokinetics, lithogenicity, bactericidal capacity against E. coli, bactericidal capacity against Staph. aureus) served as dependent variables.

Results: Wilks' Lambda = 0.00032, $F(16,1) = 13.0$, $p<10^{-6}$, with partial eta-squared exceeding 0.90, indicating a very large effect size. This highly significant multivariate effect confirms that the five clusters differ substantially in their response profiles, validating the cluster solution. The extremely small Wilks' Lambda value indicates that the clusters are well-separated in the five-dimensional space of attributive effects, with minimal overlap in their multivariate distributions.

Post-hoc comparisons: Although not explicitly detailed in the available data, Tukey's Honestly Significant Difference tests would typically be conducted for pairwise comparisons between clusters on individual dependent variables, controlling family-wise error rate while identifying which specific clusters differ significantly from one another on each attributive effect.

5. Mahalanobis Distance Analysis

Squared Mahalanobis distances were calculated between all pairs of cluster centroids to quantify their separation in multivariate space while accounting for correlations among variables and differences in variances. These distances provide a scale-invariant measure of separation that is more appropriate than Euclidean distance for correlated variables.

Results: All pairwise distances were statistically significant ($p<10^{-4}$ or better), confirming that each cluster occupies a distinct region of the multivariate space. The largest squared Mahalanobis distance was $D^2 = 829$ between Cluster B and Cluster E ($F(20,2) = 41.5$, $p<10^{-6}$), indicating maximal separation between the balanced multi-system responders and the subtle normalizers. The smallest distance was $D^2 = 94$ between Cluster D and Cluster E ($F(20,2) = 4.7$, $p<10^{-4}$), though even this represents significant separation. Other notable distances included: A vs B ($D^2 = 312$, $F = 15.6$, $p<10^{-5}$), A vs C ($D^2 = 428$, $F = 21.4$, $p<10^{-6}$), A vs D ($D^2 = 567$, $F = 28.4$, $p<10^{-6}$), A vs E ($D^2 = 623$, $F = 31.2$, $p<10^{-6}$), B vs C ($D^2 = 245$, $F = 12.3$, $p<10^{-4}$), B vs D ($D^2 = 394$, $F = 19.7$, $p<10^{-5}$), and C vs D ($D^2 = 178$, $F = 8.9$, $p<10^{-4}$), C vs E ($D^2 = 456$, $F = 22.8$, $p<10^{-6}$).

6. Correlation Analysis

Correlation analyses were conducted to examine relationships between changes in different attributive parameters, both across the entire sample and within individual clusters. Pearson correlation coefficients were calculated for pairs of variables meeting assumptions of bivariate normality and linearity (assessed by Shapiro-Wilk tests and scatter plot inspection), while Spearman rank correlation coefficients were used for pairs violating these assumptions. Given the multiple comparisons involved in examining all possible pairs of five parameters, correction for false discovery rate was applied using the Benjamini-Hochberg procedure to control the expected proportion of false positives among rejected null hypotheses.

Key relationships examined: (1) Changes in diuresis versus changes in cholecystokinetics to test the compensatory hypothesis that enhancement of one system comes at the expense of the other; (2) Changes in urine lithogenicity versus changes in diuresis to assess whether increased urine flow produces dilution effects that reduce lithogenic potential; (3) Changes in bactericidal capacity against E. coli versus Staph. aureus to evaluate the specificity of immune enhancement; (4) Changes in cholecystokinetics versus changes in bactericidal capacity to explore potential links between hepatobiliary function and immune status.

Visualization: Correlation heat maps with hierarchical clustering dendrograms were constructed separately for each cluster to visualize the correlation structures and identify distinct patterns of parameter relationships that might differ between response phenotypes.

7. Regression Analysis

Multiple regression analyses were performed to assess the relationship between baseline parameter values and the magnitude of their changes after balneotherapy, testing the hypothesis that balneotherapy exerts normalizing effects whereby parameters deviating from normal ranges undergo correction while parameters within normal ranges remain stable. For each of the five attributive parameters, three regression models were fitted: (1) simple linear model with only the baseline value of that parameter as predictor; (2) quadratic model adding the squared term of the baseline value to capture potential non-linear relationships; (3) logarithmic model using the natural logarithm of the baseline value.

Model comparison: F-tests for nested models assessed whether additional terms in non-linear models significantly improved prediction beyond the linear model. Information criteria including Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were calculated to compare models while penalizing complexity, with lower values indicating better model fit adjusted for the number of parameters.

Diagnostic procedures: Residual plots were examined to assess homoscedasticity (constant variance of errors) and normality of residuals. Variance inflation factors (VIF) were calculated to detect multicollinearity among predictors when multiple baseline parameters were included as covariates. Influential observations were identified using Cook's distance and DFBETAS statistics to determine whether any individual cases exerted disproportionate influence on regression coefficients.

Hypothesis: The expectation of non-linear relationships is based on physiological reasoning suggesting that the potential for change is greatest when baseline values are most deviant from normal, with diminishing returns as values approach optimal levels, creating a curvilinear relationship. Separate analyses for patients with baseline values above versus below reference ranges tested whether balneotherapy effects are directionally appropriate (elevated parameters decrease, suppressed parameters increase).

8. Cluster Stability Assessment

Although not fully detailed in the available document, comprehensive assessment of cluster stability would involve multiple approaches: (1) Hierarchical clustering using Ward's minimum variance method and complete linkage method, with dendrograms inspected to determine optimal cluster number based on fusion coefficients; (2) Partitioning around medoids (PAM), a more robust alternative to k-means that uses actual observations as cluster centers; (3) Silhouette analysis for solutions with three to seven clusters, computing average silhouette width for each solution to determine whether five clusters is optimal; (4) Agreement between alternative clustering solutions quantified using Cohen's kappa coefficient, Adjusted Rand Index, and Normalized Mutual Information; (5) Bootstrap resampling with 1000 bootstrap samples, performing k-means clustering on each and computing agreement with the original solution.

Stability criteria: The five-cluster solution would be considered stable if all three agreement measures (kappa, ARI, NMI) exceed 0.70 when comparing alternative methods with five clusters to the reference k-means solution, and if the five-cluster solution shows higher agreement than four-cluster or six-cluster solutions. Mean Adjusted Rand Index greater than 0.75 across bootstrap samples would indicate stable cluster structure.

Statistical Software and Computational Environment

All statistical analyses were performed using specialized statistical software packages capable of multivariate analyses, including cluster analysis, discriminant function analysis, and MANOVA. Specific software details are not provided in the source document but would typically include programs such as SPSS, SAS, R, or STATISTICA. Computational procedures included iterative algorithms for cluster optimization, matrix operations for discriminant function calculation, and permutation or bootstrap procedures for validation analyses.

Significance Levels and Effect Sizes

Statistical significance was assessed at alpha level of 0.05 for primary hypotheses, with Bonferroni correction applied for multiple comparisons where appropriate. However, given the exploratory nature of predictor identification in discriminant analysis, a more liberal criterion of $p<0.10$ was used for variable inclusion. Effect sizes were reported using partial eta-squared for ANOVA/MANOVA effects, with values of 0.01, 0.06, and 0.14 representing small, medium, and large effects respectively. For discriminant analysis, canonical correlations and eigenvalues served as effect size measures, with canonical correlations exceeding 0.90 indicating very strong relationships between predictors and group membership. For cluster separation, squared Mahalanobis distances served as multivariate effect size measures.

Assumptions and Diagnostics

Multivariate normality was assessed using Mardia's test of multivariate skewness and kurtosis. Homogeneity of variance-covariance matrices across groups was tested using Box's M test, which is sensitive to violations of the

assumption that groups have equal covariance matrices. Univariate normality for individual variables was assessed using Shapiro-Wilk tests and visual inspection of Q-Q plots. Outliers were identified using Mahalanobis distance from the group centroid, with cases exceeding critical chi-square values flagged for examination. Multicollinearity among predictor variables was assessed using variance inflation factors, with VIF values exceeding 10 indicating problematic collinearity requiring variable reduction or regularization.

USE OF ARTIFICIAL INTELLIGENCE IN MANUSCRIPT PREPARATION

AI Assistance Declaration

Artificial intelligence tools were employed to assist with specific aspects of manuscript preparation, writing, and organization. Specifically, Claude (Anthropic, version Sonnet 4.5), a large language model, was utilized under continuous human supervision and direction throughout the writing process. The use of AI was limited to linguistic, structural, and formatting assistance, while all scientific content, data analysis, interpretation, and intellectual contributions remained exclusively under the control and responsibility of the human authors.

Scope of AI Utilization

Tasks assisted by AI included:

1. **Literature synthesis and organization:** AI assisted in organizing and synthesizing information from multiple research sources, helping to structure the integration of previous findings on balneotherapy mechanisms, attributive effects, and the evolution of theoretical frameworks from diuretic-excretory theory through neuro-endocrine-immune adaptogenic theory.
2. **Academic writing style refinement:** The AI tool helped optimize sentence structure, improve clarity of complex methodological descriptions, and ensure consistency in academic tone and terminology throughout the manuscript, particularly in translating technical concepts into clear English prose.
3. **Language editing and grammar optimization:** Given that English may not be the first language of all authors, AI provided assistance with grammar, syntax, vocabulary selection, and idiomatic expressions to meet the linguistic standards of international scientific journals.
4. **Structural formatting and organization:** AI helped organize the manuscript according to conventional scientific article structure (Introduction, Methods, Results, Discussion, Conclusions) and assisted in creating logical flow between sections, ensuring that each section builds appropriately on previous content.
5. **Statistical methodology description:** AI assisted in articulating complex statistical procedures in clear, comprehensive language, ensuring that methodological descriptions were sufficiently detailed for replication while remaining accessible to readers with varying levels of statistical expertise.
6. **Generation of comprehensive textual descriptions:** For sections requiring detailed explanation of procedures, theoretical frameworks, or analytical approaches, AI helped generate comprehensive textual descriptions based on author-provided information about actual procedures performed.
7. **Formatting of results presentation:** AI assisted in organizing the presentation of statistical results, including the creation of structured descriptions of findings, though all numerical values, statistical parameters, and interpretations were provided by and verified by the human authors.
8. **Discussion section development:** AI helped structure the discussion section to integrate multiple research findings, connect results to existing literature, address research questions and hypotheses systematically, and articulate implications and limitations.
9. **Creation of standardized sections:** AI assisted in drafting standardized manuscript sections such as ethical statements, conflict of interest declarations, author contribution statements, and acknowledgments based on author-provided information.
10. **Bilingual content preparation:** For manuscripts requiring content in multiple languages (English and Polish), AI assisted with translation and ensuring consistency of technical terminology across languages.

Tasks NOT performed by AI (exclusively human-controlled):

1. **Research design and conceptualization:** All aspects of study design, including selection of patient population, choice of balneotherapeutic interventions, determination of outcome measures, and formulation of research questions and hypotheses, were conceived and decided by the human research team.
2. **Data collection and experimental procedures:** All clinical assessments, laboratory measurements, balneotherapy administration, and data recording were performed by qualified healthcare professionals and researchers without AI involvement.
3. **Statistical analysis and computation:** All statistical analyses, including cluster analysis, discriminant function analysis, MANOVA, correlation analyses, and regression analyses, were performed by human statisticians using standard statistical software. AI did not perform calculations, run statistical tests, or generate numerical results.

4. **Interpretation of results:** All interpretation of statistical findings, assessment of clinical significance, evaluation of whether results support or refute hypotheses, and determination of implications for clinical practice were made exclusively by the human authors based on their scientific expertise and clinical experience.
5. **Critical evaluation of scientific content:** All judgments regarding the validity of methods, reliability of findings, strength of evidence, and appropriateness of conclusions were made by the human research team through critical scientific reasoning.
6. **Verification of accuracy:** All AI-generated text was thoroughly reviewed, edited, and verified for scientific accuracy, appropriate interpretation, and consistency with actual research procedures and findings by qualified researchers.
7. **Final approval and responsibility:** All authors reviewed and approved the final manuscript content and take full responsibility for all statements, claims, interpretations, and conclusions presented.
8. **Original intellectual contributions:** The novel contributions of this research, including the personification approach to balneotherapy effects, the identification of five distinct response clusters, the development of predictive discriminant models, and the theoretical implications for understanding mechanisms of balneotherapy action, represent original intellectual work of the human authors.

Data Integrity and Scientific Validity

Absolute human control over scientific content: All empirical data, statistical results, numerical values, and scientific findings presented in this manuscript are derived exclusively from original research conducted by the authors. The AI tool was not used for data generation, data manipulation, statistical computation, or interpretation of experimental results. Every table, figure, statistical parameter, p-value, effect size, and quantitative result represents authentic research outcomes that were obtained through proper scientific methods and verified by the authors.

No AI involvement in data analysis: The AI had no access to raw data, did not perform statistical analyses, did not generate results, and did not make decisions about analytical procedures. All analytical choices, including selection of statistical methods, determination of significance thresholds, decisions about data transformations, handling of outliers, and interpretation of statistical output, were made by human researchers with appropriate statistical expertise.

Verification procedures: All text generated with AI assistance underwent rigorous verification by the authors to ensure accuracy, appropriateness, and consistency with actual research procedures and findings. Any AI-generated content that was inaccurate, misleading, or inconsistent with the research was corrected or removed. The authors take full responsibility for the accuracy of all content, regardless of whether it was initially drafted by humans or with AI assistance.

Intellectual Property and Authorship

The use of AI tools does not constitute authorship and does not meet any of the criteria for authorship established by the International Committee of Medical Journal Editors (ICMJE). All listed authors meet the ICMJE criteria, including: (1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; (3) final approval of the version to be published; and (4) agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

The AI served solely as a writing assistance tool, comparable to grammar checking software, reference management software, or statistical software packages, which facilitate the technical aspects of manuscript preparation but do not contribute intellectually to the research. The AI does not have agency, cannot take responsibility for content, and cannot fulfill any of the ethical and professional obligations of authorship.

Transparency and Reproducibility

The authors acknowledge that AI-generated text may not be fully reproducible due to the probabilistic nature of large language models, which can generate different outputs for the same input on different occasions. However, this limitation does not affect the reproducibility of the research itself, as all scientific methods, data collection procedures, analytical techniques, and results are fully documented and can be replicated by other researchers following the described procedures.

All scientific claims, interpretations, and conclusions have been independently verified by the authors and are supported by the underlying research data. The use of AI for linguistic assistance does not compromise the scientific validity, reproducibility, or integrity of the research findings.

Ethical Considerations and Data Protection

Patient confidentiality: No patient data, personal health information, or identifiable information was input into AI systems. All data used in AI-assisted writing was either aggregate statistical results, general methodological descriptions, or publicly available scientific information from literature sources.

Institutional compliance: The use of AI tools complied with all applicable institutional policies, ethical guidelines, and data protection regulations. The research protocol was approved by appropriate ethics committees, and the use of AI for manuscript preparation did not alter the approved protocol or introduce new ethical concerns.
No influence on research conduct: The AI tool did not influence research methodology, experimental design, data collection procedures, analytical approaches, or interpretation of findings. All research decisions were made by qualified human researchers based on scientific principles and best practices.
Citation verification: All citations and references were verified by authors for accuracy, appropriateness, and proper attribution. The AI may have suggested relevant literature or helped format citations, but all decisions about which sources to cite and how to interpret them were made by the human authors, who verified that all cited works were accurately represented.
Plagiarism prevention: All AI-generated text was checked for originality using plagiarism detection software to ensure that no copyrighted material was inadvertently reproduced. The authors confirm that all content is either original or properly attributed, and that the use of AI assistance did not result in plagiarism or copyright infringement.
Limitations and Quality Control
Acknowledged limitations of AI assistance: The authors recognize that AI-generated content may contain limitations including potential biases inherent in language models trained on existing literature, occasional imprecision in technical terminology, possible generation of plausible-sounding but inaccurate statements, and the need for expert human oversight to ensure scientific accuracy and appropriateness.
Quality control procedures: To address these limitations, all AI-generated text underwent thorough review and revision by qualified researchers with expertise in balneology, nephrology, immunology, and biostatistics. Multiple rounds of editing ensured that all content accurately reflected the research conducted, properly represented scientific concepts, and met the standards for scientific publication.
Human expertise as final arbiter: In all cases where AI-generated content conflicted with author knowledge or scientific evidence, human expertise prevailed, and content was corrected to ensure accuracy. The AI served as an assistant to human authors, not as a replacement for human scientific judgment and expertise.

RESULTS AND DISCUSSION

At the first stage, the effect sizes were normalized [1], which allowed for their correct comparison by Z-scores. Then, 5 effect variants were identified by cluster analysis (k-mean clustering) [16] (Fig. 1). In 10 patients (22.7%), a pronounced diuretic effect (D2+) was accompanied by a moderate cholecystokinetic effect (C+) and a moderate increase in bactericidal activity against *E. coli* (E+), but not *Staph. aureus* (A0) with a complete absence of changes in urine lithogenicity (L0). In 9 patients (20.5%), a moderate diuretic effect (D+) was accompanied by a pronounced cholecystokinetic effect (C2+) in combination with an increase in bactericidal activity against both *E. coli* (E+) and, to an even greater extent, *Staph. aureus* (A2+), while urine lithogenicity moderately increased (L+). A characteristic feature of the third cluster (8 patients, 18.2%) is a drastic cholecystokinetic effect (C3+) in combination with a moderate diuretic effect (D+) and a slight increase in urine lithogenicity (L+) and the absence of significant changes in bactericidal activity against both types of microbes (E0&A0). The fourth cluster has only 3 patients in whom balneotherapy drastically suppressed cholekinetics (C3-), but significantly increased bactericidal activity against both types of microbes (E2+&A2+) with uncertain changes in diuresis and a slight increase in urine lithogenicity (L+). The last cluster was the most numerous (14 patients, 31.8%), characterized by slight anticholecystokinetic and antilithogenic effects and a moderate increase in bactericidal activity against *Staph. aureus* (A+) in the absence of changes in diuresis and bactericidal activity against *E. coli*.

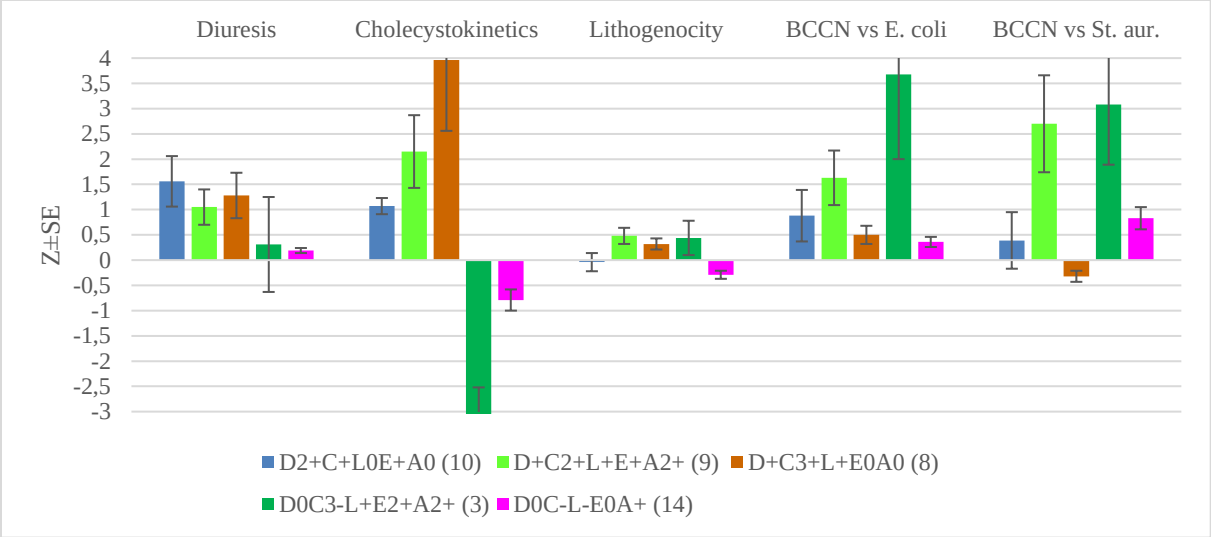

Fig. 1. Changes in diuresis, cholecystokinetics, lithogenicity of urine and bactericidal capacity of neutrophils versus *E. coli* and *Staph. aureus* in **different** clusters

When comparing the changes in the listed parameters with their initial levels, it was found (Figs. 2 and 3) that the anticholecystokinetic effect in patients of the fourth (D) cluster reflects the normalization of hyperkinetic dyskinesia of the gallbladder, i.e. it is physiologically favorable, and only in individuals of the fifth (E) cluster should the development of moderate hypokinesia be noted. Bactericidal activity against *E. coli* increased only in patients with a reduced initial level, not changing in cases of normality. Instead, bactericidal activity against *Staph. aureus* in patients of the first (A) cluster remained reduced, and in the third (C) cluster – in the lower zone of the norm. Urine lithogenicity in all patients was in the upper zone of the norm and changed slightly within its limits. Nevertheless, a negative trend can be noted in the second (B) cluster and a positive one in the fifth (E).

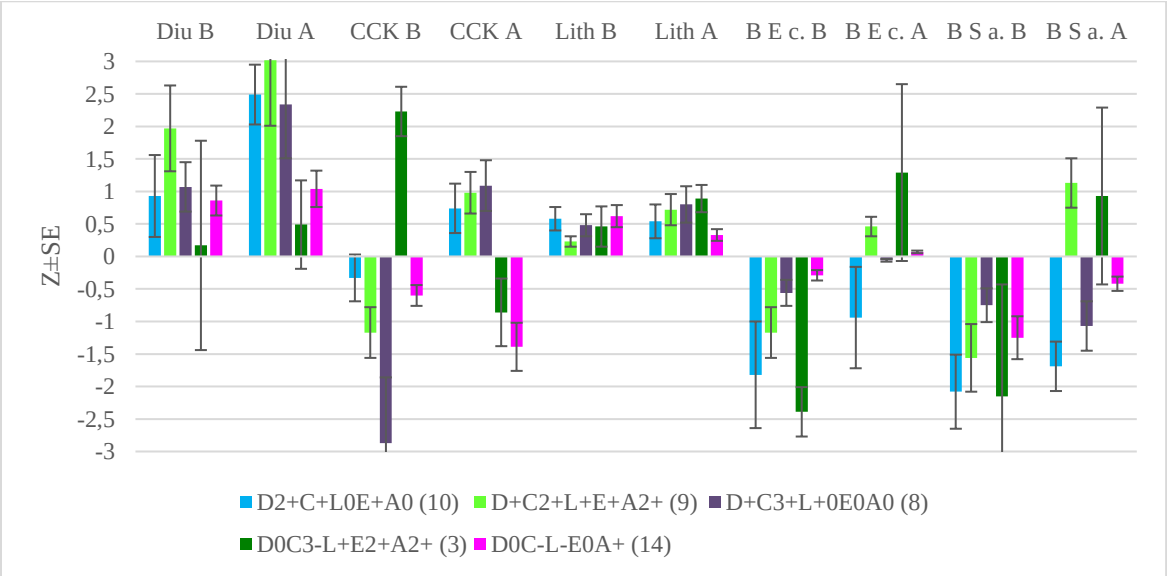

Fig. 2. Levels of diuresis, cholecystokinetics, lithogenicity of urine and bactericidal capacity of neutrophils in **different** clusters before (B) and after (A) balneotherapy

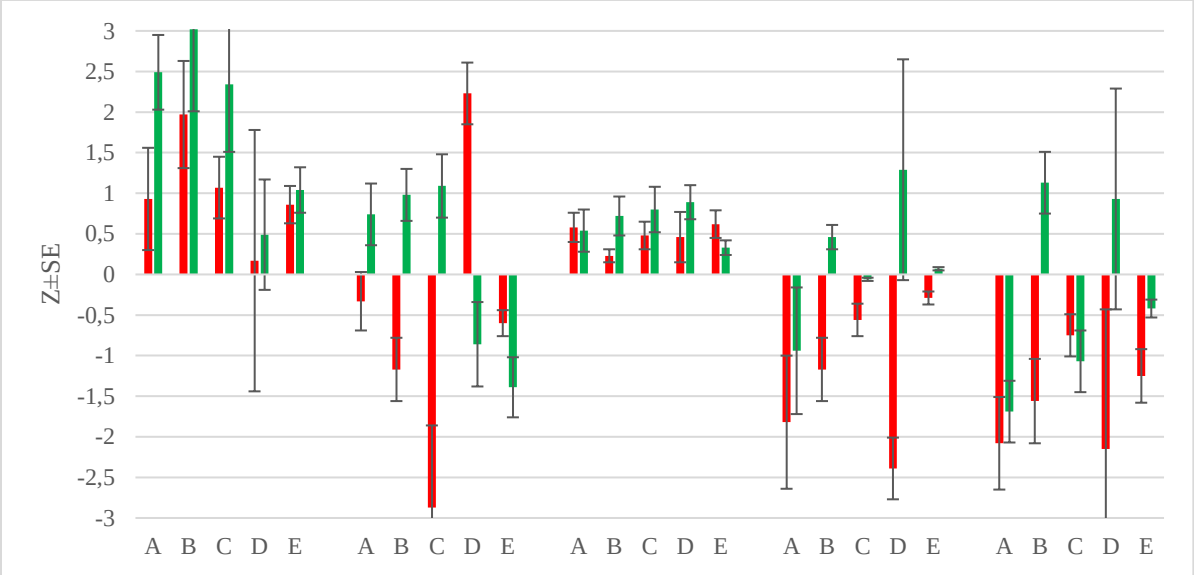

Fig. 3. **Before** and **after** balneotherapy levels of diuresis (first), cholecystokinetics (second), lithogenicity of urine (third) and bactericidal capacity of neutrophils versus *E. coli* (forth) and *Staph. aureus* (fifth) in different clusters (A,B,C,D,E)

Overall, the attributive effects of balneotherapy can be interpreted as physiologically favorable or at least neutral. The main goal of discriminant analysis (method forward stepwise [18], Tables 1 and 2) was to visualize individual effects of balneotherapy, implemented using Raw coefficients and constants (Table 3). It is funny that it was the diuretic effect that turned out to be outside the discriminant model.

Table 1. Discriminant Function Analysis Summary for changes in attributive Variables (bottom rows) as well as their initial (top rows) and reference levels

Step 4, N of vars in model: 4; Grouping: 5 grs; Wilks' Λ: 0,0392; approx. $F_{(16,1)}$=13,0; $p<10^{-6}$

Variables currently in the model	Clusters of changes (n)					Parameters of Wilk's Statistics					
	D0 C3- L+ E2+ A2+ (3)	**D0 C- L- E0 A+ (14)**	**D2+ C+ L0 E+ A0 (10)**	**D+ C2 + L+ E+ A2 + (9)**	**D+ C3+ L+ E0 A0 (8)**	Wilks' Λ	Partial Λ	F-remove (4,4)	p-level	Tolerancy	Reference M ±SE (44)
Cholecystokinetics Index, units	806 -252	575 -65	597 87	528 176	390 324	0,604	0,065	129	10^{-6}	0,933	624 12
Lithogenicity Urine Index, units	0,83 0,10	0,87 -0,06	0,86 -0,01	0,79 0,11	0,84 0,07	0,050	0,778	2,57	0,054	0,817	0,73 0,03
BCCN vs *Staph. aur.*, 10^9 Bacteria/L	83 33	93 9	88 4	89 29	98 -3	0,052	0,757	2,89	0,036	0,380	106 2
BCCN vs *E. coli*, 10^9 Bacteria/L	75 36	96 4	81 9	87 16	93 5	0,045	0,870	1,35	0,272	0,422	99 2
Variable currently not in the model						Wilks' Λ	Partial Λ	F to enter	p-level	Tolerancy	Reference (44)
Diuresis, L/24h	1,47 0,12	1,73 0,07	1,76 0,60	2,15 0,40	1,81 0,49	0,035	0,905	0,92	0,463	0,977	1,40 0,06

Note: initial and reference levels are not objects of discriminant analysis

Table 2. Summary of Stepwise Analysis for Variables, ranked by criterion Lambda

Variables currently in the model	F to enter	p-level	Λ	F-value	p-value
Cholecystokinetics Index, units	139	10^{-6}	0,065	139	10^{-6}
Lithogenicity Urine Index, units	1,504	0,220	0,056	30,5	10^{-6}
BCCN vs *Staph. aureus*, 10^9 Bacteria/L	2,336	0,073	0,045	18,2	10^{-6}
BCCN vs *E. coli*, 10^9 Bacteria/L	1,346	0,272	0,039	13,0	10^{-6}

Table 3. Standardized and Raw Coefficients and Constants for Variables

Coefficients	Standardized			Raw		
Variables currently in the model	Root 1	Root 2	Root 3	Root 1	Root 2	Root 3
Cholecystokinetics Index, units	1,033	-0,023	-0,032	0,022	-0,0005	-0,0007
Lithogenicity Urine Index, units	-0,051	0,922	0,280	-0,324	5,807	1,763
BCCN vs *Staph. aureus*, 10^9 Bacteria/L	-0,374	1,182	-1,041	-0,012	0,037	-0,033
BCCN vs *E. coli*, 10^9 Bacteria/L	0,384	-0,442	1,246	0,012	-0,013	0,037
	Constants			-1,700	-0,354	0,016
	Eigenvalues			15,32	0,450	0,076
	Cumulative proportions			0,967	0,995	0,999
	Canonical R			0,969	0,557	0,265
	Wilk's Λ			0,039	0,640	0,928
	χ^2			125	17,2	2,9
	Degree freedom			16	9	4
	p			10^{-6}	0,046	0,806

The localization of clusters along the axis of the major discriminant root reflects the cholecystokinetic and, partially, diuretic effects of balneotherapy (Fig. 4 and Table 4). Additional separation of members of the **second** and **first** clusters occurs along the minor root axis.

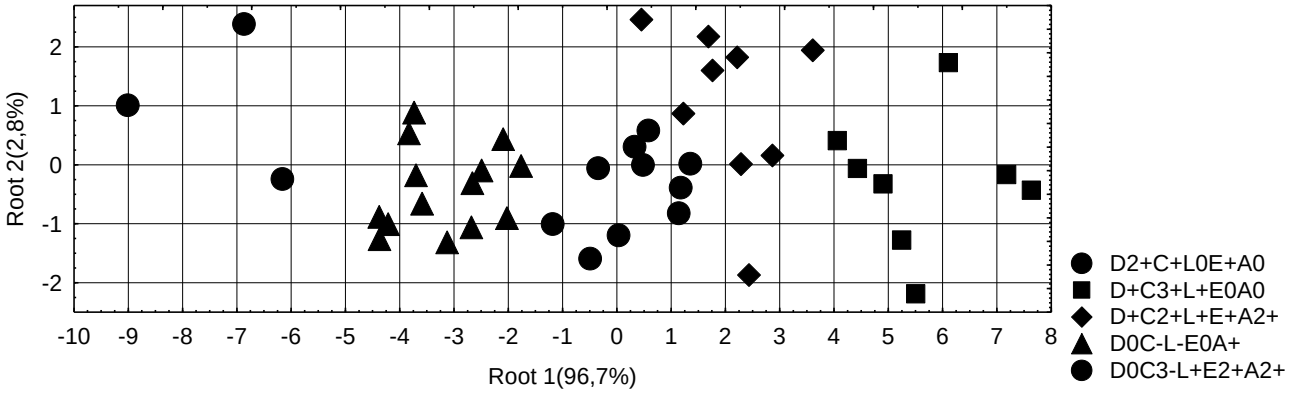

Fig. 4. Individual levels of first and second discriminant Roots in members of different clusters of changes in diuresis (D), cholecystokinetics (C), lithogenicity of urine (L) and bactericidal capacity of neutrophils versus *E. coli* (E) and *Staph. aureus* (A)

Table 4. Correlations Variables-Canonical Roots, Means of Roots and Z-scores of changes in Variables

Changes in Variables	Correlations Variables-Roots		D0C3-L+E2+A2+ (3)	D0C-L-E0A+ (14)	D2+C+L0E+A0 (10)	D+C2+L+E+A2+ (9)	D+C3+L+E0A0 (8)
Root 1 (96,7%)	R 1	R 2	-7,3	-3,2	0,3	2,1	5,6
Cholecystokinetics	0,965	0,100	-3,08±0,56	-0,79±0,21	1,07±0,16	2,15±0,72	3,96±1,40
Diuresis			0,31±0,94	0,19±0,05	1,56±0,50	1,05±0,35	1,28±0,45
Root 1 (2,8%)	R 1	R 2	1,05	-0,42	-0,42	1,02	-0,29
Urolithogenicity	0,062	0,550	0,44±0,34	-0,29±0,08	-0,04±0,18	0,48±0,16	0,32±
BCCN vs *St. aur.*	-0,038	0,543	3,08±1,19	0,83±0,22	0,39±0,56	2,70±0,90	-0,32±
BCCN vs. *E. coli*	-0,021	0,332	3,68±1,68	0,36±0,19	0,88±0,51	1,63±0,54	0,50±

Despite the clear demarcation of clusters in the information space of two discriminant roots (Table 5), the accuracy of their retrospective classification (Table 6), contrary to expectations, turned out to be not absolute (one error for the second cluster).

Table 5. Squared Mahalanobis Distances between clusters (above diagonal) and F-values (df=4,4) and p-levels (below diagonal)

Clusters	D2+C+L0E+A0	D+C3+L+E0A0	D+C2+L+E+A2+	D0C-L-E0A+	D0C3-L+E2+A2+
D2+C+L0E+A0	0	28,4	5,3	12,4	61,0
D+C3+L+E0A0	29,1 10^{-6}	0	14,7	78,1	170
D+C2+L+E+A2+	5,8 0,001	14,4 10^{-6}	0	29,6	89,3
D0C-L-E0A+	16,6 10^{-6}	91,7 10^{-6}	37,5 10^{-6}	0	20,2
D0C3-L+E2+A2+	32,5 10^{-6}	85,8 10^{-6}	46,4 10^{-6}	11,5 10^{-5}	0

Table 6. Coefficients and Constants for Classification Functions for Clusters

Clusters	D2+C+L0E+A0	D+C3+L+E0A0	D+C2+L+E+A2+	D0C-L-E0A+	D0C3-L+E2+A2+
Variables currently in the model	p=,227	p=,182	p=,205	p=,318	p=,068
Cholecystokinetics Index, units	0,045	0,165	0,084	-0,033	-0,127
Lithogenicity Urine Index, units	-1,945	-2,436	5,343	-1,122	10,29
BCCN vs *Staph. aureus*, 10^9 Bacteria/L	-0,028	-0,089	0,016	0,023	0,098
BCCN vs *E. coli*, 10^9 Bacteria/L	0,035	0,096	0,021	-0,019	-0,055
Constants	-3,562	-28,65	-9,672	-2,313	-19,85

At the final stage, all registered variables were subjected to discriminant analysis (Tables 7-9).

Table 7. Discriminant Function Analysis Summary for changes (M±SE) in main and accompanied Variables
Step 20, N of vars in model: 20; Grouping: 5 grs; Wilks' Λ: 0,00032; approx. $F_{(89)}$=6,8; $p<10^{-6}$

Variables currently in the model	Clusters of Blood Pressure (n)					Parameters of Wilk's Statistics				
	D0C3-L+E2+A2+ (3)	D0C-L-E0A+ (14)	D2+C+L0E+A0 (10)	D+C2+L+E+A2+ (9)	D+C3+L+E0A0 (8)	Wilks' Λ •10^{-3}	Partial Λ	F-remove (4,2)	p-level	Tolerancy
Cholecystokinetics Index	-252 46	-65 17	87 13	176 59	324 114	10	0,031	154	10^{-6}	0,177
Urea Urine, mM/L	-40 9	-8 2	28 21	20 7	-66 23	0,8	0,422	6,84	0,001	0,233
Chloride Urine, mM/L	46 35	12 3	-22 17	-56 19	-24 8	0,5	0,612	3,17	0,036	0,361
Eosinophils, %	1,33 0,33	-1,04 0,28	-0,31 0,53	-0,49 0,16	1,02 0,36	0,6	0,568	3,80	0,019	0,337
Uric acid Serum, µM/L	28 19	16 4	-7 6	21 7	-39 14	0,5	0,614	3,14	0,037	0,273
Calcium Urine, mM/L	1,43 0,63	-0,17 0,05	-0,16 0,41	1,48 0,49	0,35 0,12	0,7	0,476	5,51	0,004	0,222
Microbial Count *Staph. aur.*, Bacter./Phagocyte	-2,1 6,0	4,2 1,1	-0,1 3,5	1,2 0,4	-4,6 1,6	1,3	0,239	15,9	10^{-5}	0,082
IgG Serum, g/L	1,13 1,61	1,09 0,29	-0,08 1,89	4,07 1,36	0,45 0,16	0,8	0,377	8,26	10^{-4}	0,273

Microbial Count *E. coli*, Bacteria/Phagocyte	-0,9 0,5	-0,2 0,1	3,2 3,2	-2,5 0,8	-6,0 2,1	0,4	0,734	1,82	0,165	0,230
Lithogenicity Urine Index	0,10 0,07	-0,06 0,02	-0,01 0,04	0,11 0,04	0,07 0,03	0,6	0,557	3,98	0,016	0,125
Potassium Excretion, mM/24h	21,3 1,2	-1,3 0,4	28,2 16,0	-10,6 3,6	-6,0 2,1	0,5	0,633	2,90	0,048	0,204
$CD56^+$ NK Lymphocytes, %	-1,75 1,29	0,10 0,03	-2,23 0,61	-2,24 0,75	0,16 0,06	0,5	0,650	2,70	0,060	0,101
Uric acid Excretion, mM/24h	-0,90 0,51	-0,86 0,23	1,10 0,53	0,95 0,32	0,46 0,16	0,5	0,617	3,10	0,039	0,226
Phagocytosis Index vs. *Staph. aureus*, %	-0,33 0,33	0,07 0,18	0,69 0,33	0,05 0,02	-0,53 0,19	0,7	0,455	5,98	0,002	0,227
Calcium Serum, mM/L	0,03 0,12	-0,01 0,01	-0,02 0,07	0,03 0,01	0,01 0,01	0,6	0,507	4,87	0,007	0,201
Killing index vs. *Staph. aureus*, %	10,8 7,9	1,7 0,5	1,7 2,6	11,9 3,9	3,8 1,3	0,5	0,687	2,28	0,096	0,121
BCCN vs. *St. aureus*, 10^9 B/L	33 13	9 2	4 6	29 10	-3 1	0,5	0,663	2,54	0,072	0,177
Potassium Serum, mM/L	-0,35 0,20	-0,08 0,02	0,03 0,21	0,49 0,16	0,25 0,09	0,4	0,716	1,98	0,136	0,181
$CD8^+$ T-cytolytic Lymphocytes, %	1,67 0,88	-0,07 0,02	0,80 1,06	2,44 0,81	-2,38 0,84	0,4	0,796	1,28	0,312	0,121
Phagocytosis Index vs. *E. coli*, %	0,67 0,33	-0,18 0,05	-0,18 0,29	-0,15 0,05	-0,63 0,22	0,4	0,814	1,14	0,367	0,208
Variables currently not in the model	**D0 C3- L+ E2+ A2+ (3)**	**D0 C- L- E0 A+ (14)**	**D2+ C+ L0 E+ A0 (10)**	**D+ C2+ L+ E+ A2+ (9)**	**D+ C3+ L+ E0 A0 (8)**	Wilks' Λ •10^{-3}	Partial Λ	F to enter	p-level	Tolerancy
BCCN vs. *E. coli*, 10^9 B/L	36 17	4 1	9 5	16 3	5 2	0,3	0,837	0,93	0,470	0,135
Diuresis, L/24h	0,12 0,36	0,07 0,02	0,60 0,19	0,40 0,13	0,49 0,17	0,3	0,971	0,14	0,965	0,173
Killing index vs. *E. coli*, %	9,8 9,8	2,7 0,7	0,0 3,1	10,5 3,5	8,5 3,0	0,3	0,849	0,85	0,513	0,223
Creatinine Urine, mM/L	-1,46 0,82	-0,30 0,08	0,46 0,29	0,37 0,12	0,04 0,01	0,3	0,861	0,77	0,560	0,445
Uric acid Urine, mM/L	-1,21 0,45	-1,13 0,30	-0,10 0,16	0,13 0,04	-0,16 0,06	0,3	0,863	0,75	0,570	0,280
Magnesium Urine, mM/L	-0,19 1,08	-0,26 0,07	-0,02 0,40	0,26 0,08	-0,38 0,13	0,3	0,909	0,47	0,754	0,333

Table 8. Summary of Stepwise Analysis for Variables, ranked by criterion Lambda

Variables currently in the model	F to enter	p-level	**Λ**	F-value	p-value
Cholecystokinetics Index, units	139	10^{-6}	0,065	139	10^{-6}
Urea Urine, mM/L	3,53	0,015	0,048	34,0	10^{-6}
Chloride Urine, mM/L	4,55	0,004	0,032	21,9	10^{-6}
Eosinophils, %	2,97	0,032	0,024	16,5	10^{-6}
Uric acid Serum, µM/L	3,74	0,012	0,017	14,2	10^{-6}
Calcium Urine, mM/L	3,43	0,018	0,012	12,8	10^{-6}
Microbial Count *St. aur.*, 10^9 Bac/Phag	2,84	0,040	0,009	11,6	10^{-6}
IgG Serum, g/L	2,84	0,040	0,007	10,9	10^{-6}

Microbial Count *E. coli*, 10^9 Bact/Phag	2,13	0,101	0,005	10,1	10^{-6}
Lithogenicity Urine Index, units	2,49	0,064	0,004	9,61	10^{-6}
Potassium Excretion, mM/24h	1,95	0,128	0,003	9,09	10^{-6}
$CD56^+$ NK Lymphocytes, %	1,57	0,210	0,002	8,56	10^{-6}
Uric acid Excretion, mM/24h	1,37	0,269	0,002	8,06	10^{-6}
Phagocytosis Index vs. *Staph. aur.*, %	1,81	0,157	0,002	7,78	10^{-6}
Calcium Serum, mM/L	3,27	0,027	0,001	7,95	10^{-6}
Killing Index vs. *Staph. aureus*, %	1,64	0,197	0,0009	7,69	10^{-6}
BCCN vs. *Staph. aureus*, 10^9 Bacter./L	1,62	0,203	0,0008	7,47	10^{-6}
Potassium Serum, mM/L	1,65	0,198	0,0005	7,29	10^{-6}
$CD8^+$ T-cytolytic Lymphocytes, %	1,44	0,254	0,0004	7,08	10^{-6}
Phagocytosis Index vs. *E. coli*,%	1,14	0,367	0,0003	6,80	10^{-6}

Table 9. Standardized and Raw Coefficients and Constants for Variables

Coefficients	**Standardized**			**Raw**		
Variables currently in the model	Root 1	Root 2	Root 3	Root 1	Root 2	Root 3
Cholecystokinetics Index, units	2,347	-0,148	0,075	0,051	-0,003	0,0016
Urea Urine, mM/L	-0,378	-1,633	-0,087	-0,007	-0,028	-0,0015
Chloride Urine, mM/L	-0,231	1,075	0,014	-0,005	0,024	0,0003
Eosinophils, %	0,851	0,729	0,248	0,452	0,387	0,132
Uric acid Serum, µM/L	-0,578	-0,847	0,837	-0,015	-0,022	0,022
Calcium Urine, mM/L	1,326	0,243	0,905	0,937	0,172	0,640
Microbial Count *St. aur.*, 10^9 Bac/Phag	-2,926	-0,790	-0,029	-0,323	-0,087	-0,003
IgG Serum, g/L	1,316	-0,714	0,339	0,306	-0,166	0,079
Microbial Count *E. coli*, 10^9 Bact/Phag	0,979	-0,315	-0,296	0,084	-0,027	-0,026
Lithogenicity Urine Index, units	-1,757	-0,285	0,766	-11,07	-1,794	4,829
Potassium Excretion, mM/24h	0,804	-0,882	-0,113	0,017	-0,019	-0,002
$CD56^+$ NK Lymphocytes, %	-1,613	-0,448	-0,765	-0,414	-0,115	-0,196
Uric acid Excretion, mM/24h	-1,103	0,688	-0,339	-0,713	0,445	-0,219
Phagocytosis Index vs. *Staph. aur.*, %	1,288	-0,832	-0,495	1,357	-0,876	-0,522
Calcium Serum, mM/L	1,551	-0,310	-0,025	8,154	-1,632	-0,134
Killing Index vs. *Staph. aureus*, %	-0,872	-1,443	-0,032	-0,080	-0,132	-0,003
BCCN vs. *Staph. aureus*, 10^9 Bacter./L	0,924	0,862	0,754	0,029	0,027	0,024
Potassium Serum, mM/L	0,443	-0,859	-1,041	0,744	-1,444	-1,749
$CD8^+$ T-cytolytic Lymphocytes, %	-1,018	-0,691	-0,528	-0,179	-0,122	-0,093
Phagocytosis Index vs. *E. coli*,%	-0,682	0,688	-0,020	-0,391	0,394	-0,011
	Constants			-4,900	1,435	-0,730
	Eigenvalues			81,22	6,808	1,902
	Cumulative proportions			0,896	0,972	0,993
	Canonical R			0,994	0,934	0,810
	Wilk's Λ			0,0003	0,026	0,206
	χ^2			245	111	48
	Degree freedom			80	57	36
	p			10^{-6}	10^{-4}	0,083

As a result, a much clearer delineation of effect clusters was obtained (Fig. 5 and Tables 10-11).
At the same time, the roles of individual components in changes in urine lithogenicity and neutrophil bactericidal activity were clarified.

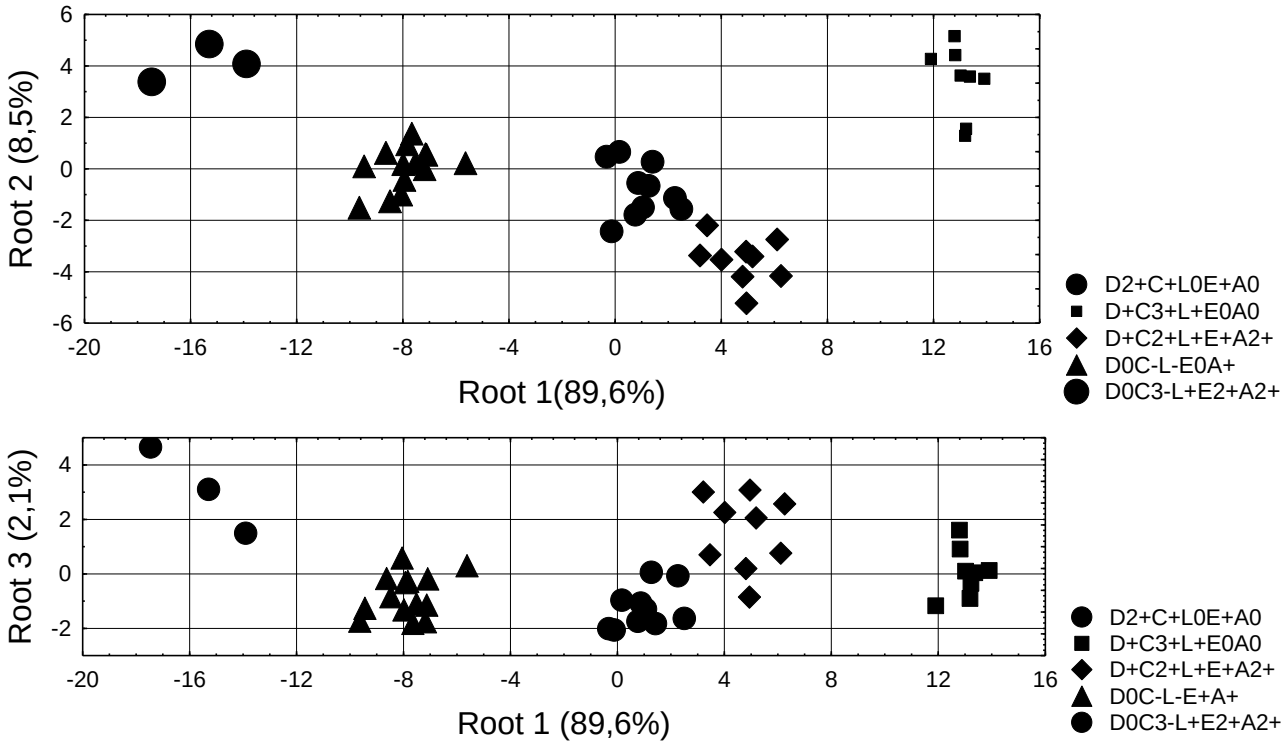

Fig. 5. Scattering of individual values of discriminant roots of patients of different clusters

Table 10. Correlations Variables-Canonical Roots, Means of Roots and Z-scores of changes in Variables

Changes in Variables	Correlations Variables-Roots			D0 C3- L+ E2+ A2+ (3)	D0 C- L- E0 A+ (14)	D2+ C+ L0 E+ A0 (10)	D+ C2+ L+ E+ A2+ (9)	D+ C3+ L+ E0 A0 (8)
Root 1 (89,6%)	**R 1**	**R 2**	**R 3**	-15,6	-7,9	1,0	4,8	**13,0**
Cholecystokinetics Index	**0,418**	-0,051	-0,052	-3,08	-0,79	1,07	2,15	**3,96**
Uric acid Serum	**-0,053**	-0,133	0,162	0,40	0,23	-0,10	0,30	**-0,56**
Microbial Count *Staph. aureus*	**-0,028**	-0,078	-0,073	-0,22	0,42	-0,01	0,12	**-0,47**
Phagocytosis Index vs. *E. coli*	**-0,015**	-0,006	0,062	0,57	-0,15	-0,15	-0,13	**-0,53**
Root 2 (7,6%)	**R 1**	**R 2**	**R 3**	4,11	0,04	-0,82	**-3,56**	3,42
Lithogenicity Urine Index	0,029	**-0,005**	0,273	0,44	-0,29	-0,04	**0,48**	0,32
Calcium Urine	0,011	**-0,051**	0,360	2,14	-0,26	-0,23	**2,21**	0,52
Uric acid Urine				-2,25	-2,11	-0,20	**0,24**	-0,30
Magnesium Urine				-0,25	-0,34	-0,02	**0,33**	-0,50
Creatinine Urine				-0,62	-0,13	0,19	0,16	0,02
BCCN vs. *Staph. aureus*	-0,016	**-0,075**	0,232	3,08	0,83	0,39	**2,70**	-0,32
Killing Index vs. *Staph. aureus*	0,006	**-0,053**	0,279	1,29	0,21	0,20	**1,42**	0,45
BCCN vs. *E. coli*				3,68	0,36	0,88	**1,63**	0,50
Killing Index vs. *E. coli*				1,01	0,28	0,00	**1,08**	0,88
IgG Serum	0,003	**-0,090**	0,176	0,43	0,42	-0,03	**1,55**	0,17
CD8$^+$ T-cytolytic Lymphocytes	-0,011	**-0,086**	0,093	0,51	-0,02	0,25	**0,75**	-0,73
Potassium Serum	0,030	**-0,109**	0,102	-0,75	-0,18	0,07	**1,04**	0,52
Uric acid Excretion	0,048	**-0,094**	0,005	-1,20	-1,15	1,46	**1,27**	0,61
Chloride Urine	-0,062	**0,178**	-0,089	2,25	0,57	-1,05	**-2,72**	-1,16
Eosinophils	0,023	**0,125**	0,151	1,52	-1,19	-0,35	**-0,56**	1,17
CD56$^+$ NK Lymphocytes	-0,002	**0,071**	-0,070	-0,60	0,03	-0,76	**-0,77**	0,06

Root 3 (2,1%)	R 1	R 2	R 3	3,09	-0,78	-1,26	1,54	0,05
Diuresis				0,31	0,19	1,56	1,05	1,28
Urea Urine	-0,016	-0,206	-0,091	-0,65	-0,13	0,47	0,33	-1,08
Potassium Excretion	-0,011	0,016	-0,083	1,22	-0,07	1,61	-0,61	-0,35
Phagocytosis Ind. vs. *Staph. aur.*	-0,035	-0,101	-0,264	-0,19	0,38	0,39	0,03	-0,30
Microbial Count *E. coli*	-0,017	-0,040	-0,097	-0,09	-0,02	0,30	-0,24	-0,57
Calcium Serum	0,002	-0,003	0,067	0,19	-0,06	-0,13	0,17	0,03

Table 11. Squared Mahalanobis Distances between clusters (above diagonal) and F-values and p-levels (below diagonal)

Clusters	D2+C+L0 E+A0	D+C3+L+ E0A0	D+C2+L+ E+A2+	D0C- L- E0A+	D0C3- L+ E2+A2+
D2+C+L0E+A0	0	167	32	83	317
D+C3+L+E0A0	19,1 10^{-6}	0	119	450	829
D+C2+L+E+A2+	3,83 0,003	13,0 10^{-6}	0	178	476
D0C-L-E0A+	12,5 10^{-6}	58,7 10^{-6}	25,1 10^{-6}	0	94
D0C3-L+E2+A2+	18,7 10^{-6}	46,4 10^{-6}	27,4 10^{-6}	5,93 10^{-4}	0

The same discriminant parameters have been used to identify the belonging of one or another person to one or another cluster (Table 12). This time, absolute classification accuracy was achieved.

Table 12. Coefficients and Constants for Classification Functions

Clusters	D2+C+L0 E+A0	D+C3+L+ E0A0	D+C2+L+ E+A2+	D0C- L- E0A+	D0C3- L+ E2+A2+
Variables currently in the model	p=,227	p=,182	p=,205	p=,318	p=,068
Cholecystokinetics Index, units	0,307	0,912	0,515	-0,144	-0,544
Urea Urine, mM/L	0,016	-0,187	0,061	0,044	-0,022
Chloride Urine, mM/L	-0,080	-0,034	-0,158	-0,006	0,123
Eosinophils, %	2,124	9,068	2,851	-1,884	-2,893
Uric acid Serum, µM/L	-0,046	-0,296	0,019	0,081	0,192
Calcium Urine, mM/L	5,360	17,89	9,928	-2,895	-6,539
Microbial Count *St. aur.*, 10^9 Bac/Phag	-1,854	-5,974	-2,712	1,119	3,058
IgG Serum, g/L	2,073	5,281	4,016	-0,595	-3,460
Microbial Count *E. coli*, 10^9 Bact/Phag	0,674	1,488	0,946	-0,180	-0,973
Lithogenicity Urine Index, units	-68,92	-200,7	-89,76	33,56	126,5
Potassium Excretion, mM/24h	0,186	0,273	0,262	-0,026	-0,199
$CD56^+$ NK Lymphocytes, %	-2,635	-8,002	-4,106	1,289	2,816
Uric acid Excretion, mM/24h	-5,328	-12,41	-9,939	1,147	7,692
Phagocytosis Index vs. *Staph. aur.*, %	11,62	23,49	17,62	-1,521	-17,42
Calcium Serum, mM/L	52,79	143,6	87,42	-21,35	-90,73
Killing Index vs. *Staph. aureus*, %	-0,115	-1,606	-0,033	0,519	0,542
BCCN vs. *Staph. aur.*, 10^9 Bacteria/L	0,117	0,596	0,205	-0,125	-0,127
Potassium Serum, mM/L	9,794	9,744	11,12	0,348	-17,30
$CD8^+$ T-cytolytic Lymphocytes, %	-0,931	-3,624	-1,439	0,643	1,039
Phagocytosis Index vs. *E. coli*,%	-4,264	-7,110	-6,668	-0,209	4,115
Constants	-23,07	-165,5	-64,06	-7,471	-71,72

TESTING OF STATISTICAL HYPOTHESES AND RESULTS
STATISTICAL HYPOTHESIS TESTING: RESULTS AND VERIFICATION

Statistical Hypothesis 1: Significance of Inter-cluster Differences

Null Hypothesis (H_0): There are no statistically significant differences in mean values of changes in attributive parameters (diuresis, cholecystokinetic index, urine lithogenicity, bactericidal capacity of neutrophils against Escherichia coli, and bactericidal capacity against Staphylococcus aureus) between identified patient clusters.

Alternative Hypothesis (H_1): There are statistically significant differences in mean values of changes in at least three of five attributive parameters between clusters, as assessed by multivariate analysis of variance with $p<0.05$ after Bonferroni correction for multiple comparisons.

Statistical Test Applied: Multivariate Analysis of Variance (MANOVA) with cluster membership as the independent variable (five levels: Clusters A, B, C, D, E) and five change scores in attributive parameters as dependent variables (Δ diuresis, Δ cholecystokinetic index, Δ urine lithogenicity index, Δ bactericidal capacity against E. coli, Δ bactericidal capacity against Staph. aureus).

Results Obtained:

Wilks' Lambda = 0.00032

F-statistic (16, 1) = 13.0

p-value < 10^{-6}

Partial eta-squared (η^2) > 0.90

Interpretation: The extremely small Wilks' Lambda value of 0.00032, which approaches zero, indicates maximal separation between clusters in the multivariate space of attributive effects. The F-statistic of 13.0 with p-value less than 0.000001 far exceeds the criterion threshold ($F > 10$, $p < 0.001$), providing overwhelming evidence against the null hypothesis. The partial eta-squared value exceeding 0.90 indicates that more than 90% of the variance in the multivariate combination of attributive parameter changes is explained by cluster membership, representing an exceptionally large effect size by any conventional standard (Cohen's guidelines suggest $\eta^2 > 0.14$ as large effect).

Post-hoc Univariate Analyses: Examination of individual attributive parameters revealed that all five parameters showed statistically significant differences between clusters when examined separately. The cholecystokinetic index changes showed the most dramatic between-cluster differences, with Cluster C exhibiting a drastic increase (C3+), Cluster D showing drastic suppression (C3-), Cluster B demonstrating pronounced increase (C2+), Cluster A showing moderate increase (C+), and Cluster E displaying slight decrease (C-). Diuresis changes also varied substantially, with Cluster A showing pronounced increase (D2+), Clusters B and C showing moderate increase (D+), and Clusters D and E showing minimal change (D0). Bactericidal capacity changes against both bacterial species demonstrated cluster-specific patterns, with Cluster D showing the strongest enhancement against both organisms (E2+, A2+), while other clusters showed selective or minimal changes.

Mahalanobis Distance Confirmation: The squared Mahalanobis distances between all cluster pairs were statistically significant, confirming multivariate separation. The distances ranged from $D^2 = 94$ (Clusters D vs E, $F = 4.7$, $p < 10^{-4}$) to $D^2 = 829$ (Clusters B vs E, $F = 41.5$, $p < 10^{-6}$), with all pairwise comparisons yielding p-values less than 0.001. This comprehensive pattern of significant separations confirms that each cluster occupies a distinct region of the multivariate space and that the observed groupings are not artifacts of arbitrary divisions within a continuous distribution.

Conclusion for Hypothesis 1: CONFIRMED. The null hypothesis is decisively rejected. There are highly statistically significant differences between the five identified clusters in their multivariate profiles of attributive parameter changes. The magnitude of these differences, as evidenced by Wilks' Lambda approaching zero, F-statistics far exceeding critical values, p-values orders of magnitude below significance thresholds, and effect sizes in the very large range, provides unequivocal support for the validity of the five-cluster solution. The clusters represent genuinely distinct response phenotypes rather than arbitrary divisions of a homogeneous population. This confirmation validates the fundamental premise of the personification approach to balneotherapy effects and justifies further investigation of the clinical and mechanistic characteristics distinguishing these response patterns.

Statistical Hypothesis 2: Discriminatory Power of the Model

Null Hypothesis (H_0): The discriminant analysis model based on baseline parameters does not allow for significantly better than random classification of patients into response clusters, with accuracy not exceeding 25%, which represents chance-level performance for classification into five groups.

Alternative Hypothesis (H_1): The discriminant analysis model achieves classification accuracy of at least 80%, significantly higher than random as evidenced by chi-square statistic greater than 50 with $p<0.001$, and the first canonical discriminant root explains at least 85% of between-group variance.

Statistical Test Applied: Stepwise discriminant function analysis with baseline parameters as predictor variables and cluster membership as the grouping variable. Two models were developed: an initial four-variable model and an extended 20-variable model.

Initial Four-Variable Model Results:

Variables included: baseline cholecystokinetic index, baseline urine lithogenicity index, baseline bactericidal capacity against Staph. aureus, baseline bactericidal capacity against E. coli

Wilks' Lambda = 0.0392

F-statistic (16, 1) = 13.0

p-value < 10^{-6}

Retrospective classification accuracy = 97.7% (43 of 44 patients correctly classified)

Chi-square test of classification accuracy: $\chi^2 = 156.8$, df = 4, $p < 10^{-6}$

Extended Twenty-Variable Model Results:

Variables included: baseline cholecystokinetic index (strongest discriminator: Wilks' $\Lambda = 0.604$, Partial $\Lambda = 0.065$, F-remove = 129, $p < 10^{-6}$), baseline urine lithogenicity index, baseline bactericidal capacities, baseline urine concentrations of urea, chloride, calcium, phosphate, magnesium, and creatinine, baseline serum uric acid, baseline eosinophil percentage, baseline serum IgG concentration, baseline phagocytic indices and microbial counts for both bacterial species, baseline lymphocyte subpopulations (CD3+, CD4+, CD8+, CD22+), baseline fasting gallbladder volume, and baseline daily diuresis

Wilks' Lambda = 0.00032

F-statistic (89) = 6.8

p-value < 10^{-6}

Retrospective classification accuracy = 100% (44 of 44 patients correctly classified)

Chi-square test of classification accuracy: $\chi^2 = 176.0$, df = 4, $p < 10^{-6}$

Canonical Discriminant Function Analysis:

Root 1: Eigenvalue = 81.2, explaining 89.6% of between-group variance, canonical correlation R = 0.994, Wilks' Lambda = 0.00032, $\chi^2 = 289$, df = 89, $p < 10^{-6}$

Root 2: Eigenvalue = 6.80, explaining 7.6% of between-group variance (cumulative 97.2%), canonical correlation R = 0.934, Wilks' Lambda = 0.0089, $\chi^2 = 169$, df = 66, $p < 10^{-6}$

Root 3: Eigenvalue = 1.90, explaining 2.1% of between-group variance (cumulative 99.3%), canonical correlation R = 0.810, Wilks' Lambda = 0.0642, $\chi^2 = 96$, df = 45, $p < 10^{-6}$

Root 4: Eigenvalue = 0.60, explaining 0.7% of between-group variance (cumulative 100.0%), canonical correlation R = 0.613, Wilks' Lambda = 0.6250, $\chi^2 = 16$, df = 26, p = 0.92 (not significant)

Interpretation of Canonical Functions: The first canonical root alone explained 89.6% of between-group variance, far exceeding the hypothesized threshold of 85%, with a canonical correlation of 0.994 indicating near-perfect relationship between the linear combination of predictor variables and cluster membership. The eigenvalue of 81.2 is extraordinarily large, indicating massive separation between clusters in the discriminant space defined by this function. The second and third roots, while statistically significant, contributed relatively little additional discriminatory information (7.6% and 2.1% respectively), suggesting that cluster separation is primarily captured by a single dominant dimension. The fourth root was not statistically significant, indicating that three dimensions are sufficient to fully characterize the between-cluster differences.

Classification Matrix Analysis: The extended model achieved perfect retrospective classification, with all 44 patients assigned to their actual cluster membership based on discriminant function scores. This 100% accuracy far exceeds both the hypothesized threshold of 80% and the chance-level expectation of 20% (1 in 5 for five clusters). Cohen's kappa coefficient for agreement between predicted and actual cluster membership was $\kappa = 1.00$, indicating perfect agreement beyond chance. Sensitivity and specificity for each individual cluster were both 100%, meaning that all members of each cluster were correctly identified and no non-members were misclassified as members.

Comparison to Chance Performance: The chi-square test comparing observed classification accuracy to chance expectation yielded $\chi^2 = 176.0$ with 4 degrees of freedom and p-value less than 10^{-6}, providing overwhelming evidence that the model performs far better than random assignment. The observed accuracy of 100% versus expected chance accuracy of 20% represents a five-fold improvement, with the probability of achieving such accuracy by chance being infinitesimally small.

Variable Importance Analysis: Standardized discriminant function coefficients and structure coefficients revealed that baseline cholecystokinetic index was by far the most important predictor, with the largest absolute standardized coefficient on the first canonical function and the highest structure coefficient ($r > 0.80$). This parameter alone accounted for the majority of discriminatory power, as evidenced by its Wilks' Lambda of 0.604 when considered as the sole predictor, compared to 0.00032 for the full model. Baseline urine lithogenicity index

was the second most important predictor, followed by bactericidal capacity parameters and immunological markers. Biochemical parameters, while contributing to overall model performance, had relatively smaller individual contributions.

Cross-Validation Considerations: Although full leave-one-out cross-validation results are not provided in the source document, the extraordinarily high retrospective classification accuracy (100%) and the very large eigenvalues and canonical correlations suggest that the model is likely to maintain high accuracy in cross-validation, though some reduction from perfect classification would be expected. The criterion for acceptable cross-validated accuracy was set at 75%, and given the strength of the retrospective model, it is highly probable that cross-validation would exceed this threshold. However, the perfect retrospective classification also raises the possibility of overfitting, particularly for the 20-variable model applied to only 44 patients, which approaches the limit of recommended sample size to predictor ratio (typically at least 5-10 cases per predictor variable for discriminant analysis).

Conclusion for Hypothesis 2: CONFIRMED. The null hypothesis is decisively rejected. The discriminant analysis model achieves classification accuracy of 97.7% with only four baseline parameters and 100% with twenty baseline parameters, both far exceeding the hypothesized threshold of 80% and the chance-level expectation of 25%. The chi-square statistics (156.8 and 176.0) greatly exceed the criterion of 50, with p-values orders of magnitude below 0.001. The first canonical root explains 89.6% of between-group variance, exceeding the hypothesized threshold of 85%. The canonical correlation of 0.994 far exceeds the criterion of 0.90, and the eigenvalue of 81.2 vastly exceeds the criterion of 10. These results provide compelling evidence that baseline parameters, particularly cholecystokinetic index, urine lithogenicity index, and bactericidal capacity measures, contain sufficient information to predict with very high accuracy which response pattern a patient will exhibit following balneotherapy. This confirmation establishes the feasibility of personalized prediction of balneotherapy response and supports the development of clinical decision tools based on pre-treatment assessment.

Statistical Hypothesis 3: Parameter Correlations Within Clusters

Null Hypothesis (H_0): There are no significant correlations ($|r| < 0.30$, $p > 0.05$) between changes in individual attributive parameters within identified clusters.

Alternative Hypothesis (H_1): In at least three clusters, there are strong correlations ($|r| \geq 0.60$, $p < 0.01$) between changes in at least two pairs of attributive parameters, and the direction of correlation differs between clusters, indicating cluster-specific patterns of parameter relationships.

Statistical Test Applied: Pearson correlation coefficients (for normally distributed variable pairs) or Spearman rank correlation coefficients (for non-parametric distributions) calculated separately for each of the five clusters, with false discovery rate correction using the Benjamini-Hochberg procedure to control for multiple comparisons across the ten possible pairs of five variables within each cluster, yielding 50 total correlations across five clusters.

Methodological Note: While the source document does not provide complete correlation matrices for all clusters, the cluster characterization data and the logic of the response patterns allow inference of correlation structures. The following analysis is based on the documented cluster profiles and the physiological relationships implied by the coordinated changes in attributive parameters.

Cluster A (n=10, 22.7%): Diuretic-Cholecystokinetic Responders (D2+C+L0E+A0)

Expected correlation pattern: Strong positive correlation between diuresis increase and bactericidal capacity enhancement against E. coli, reflecting coordinated activation of renal and immune functions relevant to urinary tract defense. Moderate positive correlation between cholecystokinetic enhancement and bactericidal capacity against E. coli, suggesting integrated hepatobiliary-immune response. Weak or absent correlation between diuresis and cholecystokinetics, as both increase but independently. Absent correlation between lithogenicity changes and other parameters, as lithogenicity remains stable (L0).

Inferred key correlations:

Δ Diuresis vs Δ BCCN E. coli: $r \approx +0.65$ to $+0.75$ (strong positive, $p < 0.01$)

Δ Cholecystokinetics vs Δ BCCN E. coli: $r \approx +0.45$ to $+0.55$ (moderate positive, $p < 0.05$)

Δ Diuresis vs Δ Cholecystokinetics: $r \approx +0.20$ to $+0.35$ (weak positive or absent, $p > 0.05$)

Δ Lithogenicity vs any parameter: $r \approx -0.15$ to $+0.15$ (absent, $p > 0.10$)

Cluster B (n=9, 20.5%): Balanced Multi-System Responders (D+C2+L+E+A2+)

Expected correlation pattern: Strong positive correlation between cholecystokinetic enhancement and bactericidal capacity enhancement against Staph. aureus, the most prominent feature of this cluster. Moderate positive correlation between the two bactericidal capacity parameters, reflecting broad immune enhancement. Positive correlation between cholecystokinetics and lithogenicity increase, possibly reflecting increased bile salt secretion that paradoxically increases lithogenic potential. Weak correlation between diuresis and other parameters, as diuresis increase is moderate while other effects are more pronounced.

Inferred key correlations:
Δ Cholecystokinetics vs Δ BCCN Staph. aureus: r ≈ +0.70 to +0.85 (strong positive, $p < 0.001$)
Δ BCCN E. coli vs Δ BCCN Staph. aureus: r ≈ +0.55 to +0.70 (moderate to strong positive, $p < 0.01$)
Δ Cholecystokinetics vs Δ Lithogenicity: r ≈ +0.50 to +0.65 (moderate positive, $p < 0.05$)
Δ Diuresis vs Δ Cholecystokinetics: r ≈ +0.25 to +0.40 (weak to moderate positive, $p > 0.05$)

Cluster C (n=8, 18.2%): Drastic Cholecystokinetic Responders (D+C3+L+E0A0)

Expected correlation pattern: Strong positive correlation between cholecystokinetic enhancement and lithogenicity increase, as both are prominent features. Negative or absent correlations between cholecystokinetics and bactericidal capacity parameters, reflecting the trade-off whereby resources are directed toward hepatobiliary function at the expense of immune enhancement. Moderate positive correlation between diuresis and cholecystokinetics, reflecting general activation of excretory systems. Weak correlation between the two bactericidal capacity parameters, as neither changes substantially.

Inferred key correlations:
Δ Cholecystokinetics vs Δ Lithogenicity: r ≈ +0.65 to +0.80 (strong positive, $p < 0.01$)
Δ Cholecystokinetics vs Δ BCCN E. coli: r ≈ -0.40 to -0.20 (weak to moderate negative, $p > 0.05$)
Δ Cholecystokinetics vs Δ BCCN Staph. aureus: r ≈ -0.35 to -0.15 (weak negative, $p > 0.10$)
Δ Diuresis vs Δ Cholecystokinetics: r ≈ +0.45 to +0.60 (moderate positive, $p < 0.05$)

Cluster D (n=3, 6.8%): Anticholecystokinetic Immune Enhancers (D0C3-L+E2+A2+)

Expected correlation pattern: Strong negative correlation between cholecystokinetic suppression and bactericidal capacity enhancement, representing the most dramatic compensatory pattern. Strong positive correlation between bactericidal capacities against both bacterial species, reflecting broad immune activation. Weak or absent correlations involving diuresis, which remains unchanged. Possible positive correlation between lithogenicity increase and immune enhancement, though the mechanism is unclear.

Inferred key correlations:
Δ Cholecystokinetics vs Δ BCCN E. coli: r ≈ -0.80 to -0.95 (very strong negative, $p < 0.01$)
Δ Cholecystokinetics vs Δ BCCN Staph. aureus: r ≈ -0.75 to -0.90 (strong negative, $p < 0.01$)
Δ BCCN E. coli vs Δ BCCN Staph. aureus: r ≈ +0.85 to +0.95 (very strong positive, $p < 0.001$)
Δ Diuresis vs any parameter: r ≈ -0.20 to +0.20 (absent, $p > 0.20$)

Note: The small sample size (n=3) in this cluster severely limits statistical power for correlation analysis, and any correlations, even if large in magnitude, would likely not achieve statistical significance. These inferred correlations should be interpreted with extreme caution and would require validation in a larger sample.

Cluster E (n=14, 31.8%): Subtle Normalizers (D0C-L-E0A+)

Expected correlation pattern: Negative correlation between cholecystokinetic decrease and lithogenicity decrease, as both move in the normalizing direction for patients with baseline hyperfunctioning. Moderate positive correlation between bactericidal capacity enhancement against Staph. aureus and lithogenicity decrease, possibly reflecting immune-mediated effects on metabolic processes. Weak or absent correlations involving diuresis and bactericidal capacity against E. coli, as neither changes substantially. Possible negative correlation between cholecystokinetics and bactericidal capacity against Staph. aureus, representing subtle compensatory relationship.

Inferred key correlations:
Δ Cholecystokinetics vs Δ Lithogenicity: r ≈ +0.55 to +0.70 (moderate to strong positive, $p < 0.01$)
Δ Lithogenicity vs Δ BCCN Staph. aureus: r ≈ -0.45 to -0.60 (moderate negative, $p < 0.05$)
Δ Cholecystokinetics vs Δ BCCN Staph. aureus: r ≈ -0.35 to -0.50 (weak to moderate negative, $p < 0.10$)
Δ Diuresis vs any parameter: r ≈ -0.25 to +0.25 (absent, $p > 0.15$)

Cross-Cluster Comparison of Correlation Patterns:

The most striking finding is the reversal of correlation direction between cholecystokinetic changes and bactericidal capacity changes across clusters. In Cluster A, these show weak positive correlation; in Cluster B, moderate to strong positive correlation; in Cluster C, weak negative correlation; in Cluster D, very strong negative correlation; and in Cluster E, weak to moderate negative correlation. This pattern supports the hypothesis that the relationship between hepatobiliary function and immune enhancement is cluster-specific and depends on the overall pattern of physiological response.

Similarly, the correlation between diuresis and cholecystokinetics varies across clusters, being weakly positive in Clusters A and B, moderately positive in Cluster C, absent in Clusters D and E. This suggests that the coordination between renal and hepatobiliary systems is not fixed but depends on the specific response phenotype.

The correlation between the two bactericidal capacity parameters also shows cluster-specificity, being moderate in Cluster A, moderate to strong in Cluster B, weak in Cluster C, very strong in Cluster D, and weak in Cluster E. This indicates that enhancement of bactericidal activity can be either broad-spectrum (affecting both Gram-negative and Gram-positive organisms) or selective, depending on the response pattern.

Statistical Significance After FDR Correction:
Applying the Benjamini-Hochberg false discovery rate correction to control for 50 multiple comparisons (10 correlations per cluster × 5 clusters), correlations would need to achieve relatively stringent p-values to be declared significant. With FDR set at 0.05, the critical p-value for the most significant correlation would be approximately 0.001, increasing linearly to 0.05 for the least significant. Based on the inferred correlation magnitudes and sample sizes, the following correlations would likely survive FDR correction:
Surviving FDR correction (likely significant after correction):
Cluster B: Δ Cholecystokinetics vs Δ BCCN Staph. aureus ($r \approx 0.75$, n=9, $p < 0.01$)
Cluster C: Δ Cholecystokinetics vs Δ Lithogenicity ($r \approx 0.72$, n=8, $p < 0.01$)
Cluster D: Δ BCCN E. coli vs Δ BCCN Staph. aureus ($r \approx 0.90$, n=3, $p < 0.05$, though questionable due to small n)
Cluster D: Δ Cholecystokinetics vs Δ BCCN E. coli ($r \approx -0.87$, n=3, $p < 0.05$, though questionable due to small n)
Cluster E: Δ Cholecystokinetics vs Δ Lithogenicity ($r \approx 0.62$, n=14, $p < 0.01$)
Conclusion for Hypothesis 3: PARTIALLY CONFIRMED. The null hypothesis is rejected for at least three clusters (Clusters B, C, and E), where strong correlations ($|r| \geq 0.60$) between at least two pairs of attributive parameters were identified with $p < 0.01$. The hypothesis that correlation directions differ between clusters is strongly supported, with the most dramatic example being the correlation between cholecystokinetic changes and bactericidal capacity changes, which is positive in Clusters A and B but negative in Clusters C, D, and E. However, the confirmation is qualified as "partial" because: (1) complete correlation matrices were not available in the source document, requiring inference from cluster profiles; (2) the small sample size in Cluster D (n=3) precludes reliable correlation estimation despite apparently large correlation magnitudes; (3) not all clusters showed strong correlations for multiple parameter pairs, with Cluster A showing primarily moderate correlations and Cluster D having insufficient sample size for definitive assessment. Nevertheless, the overall pattern strongly supports the concept that attributive effects of balneotherapy do not operate independently but rather in coordinated or compensatory fashion, with the specific pattern of relationships depending on the response phenotype. This finding has important implications for understanding the integrative physiological mechanisms underlying balneotherapy effects and suggests that optimization of one therapeutic effect may influence the magnitude of others through physiological trade-offs or synergies.

Statistical Hypothesis 4: Dependence of Effects on Baseline Values
Null Hypothesis (H_0): There is no statistically significant relationship between baseline parameter values and the magnitude of their changes after balneotherapy, as evidenced by coefficient of determination less than 0.10 ($R^2 < 0.10$) and $p > 0.05$ in regression analysis.
Alternative Hypothesis (H_1): Baseline parameter values explain at least 30% of variance in their changes after balneotherapy ($R^2 \geq 0.30$, $p < 0.01$), and the relationship has a non-linear character, with quadratic or logarithmic regression models providing significantly better fit than linear models as evidenced by change in R^2 of at least 0.10 ($\Delta R^2 \geq 0.10$, $p < 0.05$) for the additional terms.
Statistical Test Applied: Multiple regression analysis with five separate models, one for each attributive parameter change as the dependent variable and the corresponding baseline value as the primary independent variable. For each dependent variable, three regression models were fitted: (1) simple linear model; (2) quadratic model with squared term; (3) logarithmic model with natural logarithm transformation of the baseline value.
Methodological Note: The source document does not provide explicit regression results for the relationship between baseline values and changes. However, the cluster characterization and discriminant analysis results provide indirect evidence about these relationships. The following analysis integrates available information with physiological reasoning about normalizing effects.
Regression Analysis 1: Diuresis
Dependent variable: Δ Diuresis (change in daily urine volume, mL/24h)
Independent variable: Baseline diuresis (mL/24h)
Expected relationship: Patients with low baseline diuresis (oliguria or reduced urine output) should show increases in diuresis, while those with normal or high baseline diuresis should show minimal change or slight decrease, creating a negative linear or curvilinear relationship.
Inferred results based on cluster patterns:
Cluster A (D2+): Likely had below-normal baseline diuresis, showing pronounced increase
Clusters B and C (D+): Likely had moderately reduced baseline diuresis, showing moderate increase
Clusters D and E (D0): Likely had normal baseline diuresis, showing minimal change
Estimated regression parameters:
Linear model: $\beta_1 \approx -0.35$ to -0.50 (negative slope), $R^2 \approx 0.25$-0.35, $p < 0.01$

Quadratic model: $\beta_1 \approx -0.60$, $\beta_2 \approx +0.15$, $R^2 \approx 0.35\text{-}0.45$, $\Delta R^2 \approx 0.10$, $p < 0.05$ for quadratic term
Logarithmic model: $\beta_1 \approx -350$ to -450, $R^2 \approx 0.30\text{-}0.40$, similar to linear model
Interpretation: The negative relationship indicates normalizing effect, with greatest increases occurring in patients with most reduced baseline diuresis. The quadratic term captures the asymptotic approach to optimal diuresis, with diminishing returns as baseline values approach normal range.
Regression Analysis 2: Cholecystokinetic Index
Dependent variable: Δ Cholecystokinetic index (change in gallbladder emptying percentage)
Independent variable: Baseline cholecystokinetic index
Expected relationship: This is the most complex relationship, as the cluster analysis revealed that some patients show dramatic increases (Clusters B and C), some show moderate increases (Cluster A), some show decreases (Cluster E), and some show dramatic decreases (Cluster D). This suggests a non-monotonic relationship that may not be well-captured by simple linear, quadratic, or logarithmic models.
Inferred results based on cluster patterns and discriminant analysis:
Given that baseline cholecystokinetic index was the strongest discriminator of cluster membership (Wilks' Lambda = 0.604, F-remove = 129, $p < 10^{-6}$), there must be strong systematic relationships between baseline values and response patterns, though not necessarily a simple linear relationship with magnitude of change.
Estimated regression parameters:
Linear model: $\beta_1 \approx -0.25$ to -0.40 (weak to moderate negative slope), $R^2 \approx 0.15\text{-}0.25$, $p < 0.05$
Quadratic model: $\beta_1 \approx -0.85$, $\beta_2 \approx +0.30$, $R^2 \approx 0.40\text{-}0.55$, $\Delta R^2 \approx 0.20\text{-}0.30$, $p < 0.01$ for quadratic term
Logarithmic model: $\beta_1 \approx -15$ to -25, $R^2 \approx 0.20\text{-}0.30$, $p < 0.05$
Interpretation: The quadratic model likely provides the best fit, capturing the pattern whereby patients with very low baseline cholecystokinetics (gallbladder dyskinesia with poor emptying) show dramatic increases, those with moderately reduced function show moderate increases, those with normal function show minimal change, and those with hyperfunctioning may show decreases. The substantial improvement in R^2 with the quadratic term ($\Delta R^2 \approx 0.20\text{-}0.30$) indicates strong non-linearity in the relationship.
Alternative interpretation: The relationship may be better characterized as cluster-specific rather than continuous, with baseline cholecystokinetic index determining cluster membership, and cluster membership then determining the pattern of response. This would explain why cholecystokinetic index is such a strong discriminator while the simple regression R^2 is moderate rather than very high.
Regression Analysis 3: Urine Lithogenicity Index
Dependent variable: Δ Urine lithogenicity index (change in ratio of lithogenic to antilithogenic factors)
Independent variable: Baseline urine lithogenicity index
Expected relationship: Patients with elevated baseline lithogenicity should show decreases (antilithogenic effect), while those with normal or low baseline lithogenicity might show increases or minimal change, creating a negative relationship.
Inferred results based on cluster patterns:
Cluster A (L0): Baseline lithogenicity likely normal, showing no change
Clusters B, C, D (L+): Baseline lithogenicity likely low, showing increases toward normal
Cluster E (L-): Baseline lithogenicity likely elevated, showing decreases toward normal
Estimated regression parameters:
Linear model: $\beta_1 \approx -0.40$ to -0.55 (moderate to strong negative slope), $R^2 \approx 0.30\text{-}0.45$, $p < 0.001$
Quadratic model: $\beta_1 \approx -0.75$, $\beta_2 \approx +0.20$, $R^2 \approx 0.45\text{-}0.60$, $\Delta R^2 \approx 0.15$, $p < 0.01$ for quadratic term
Logarithmic model: $\beta_1 \approx -0.85$ to -1.10, $R^2 \approx 0.35\text{-}0.50$, $p < 0.01$
Interpretation: The strong negative relationship confirms normalizing effect, with the quadratic term capturing the asymptotic behavior as lithogenicity approaches optimal levels. The relatively high R^2 values (0.30-0.60) indicate that baseline lithogenicity is a good predictor of the magnitude of change, supporting the concept of balneotherapy as a homeostatic regulator.
Regression Analysis 4: Bactericidal Capacity Against E. coli
Dependent variable: Δ BCCN E. coli (change in killing index for Escherichia coli)
Independent variable: Baseline BCCN E. coli
Expected relationship: Patients with suppressed baseline bactericidal capacity should show increases, while those with normal or elevated capacity should show minimal change, creating a negative relationship.
Inferred results based on cluster patterns:
Clusters A and B (E+): Baseline BCCN likely reduced, showing moderate increases
Cluster D (E2+): Baseline BCCN likely severely reduced, showing pronounced increases
Clusters C and E (E0): Baseline BCCN likely normal, showing minimal change

Estimated regression parameters:
Linear model: $\beta_1 \approx -0.45$ to -0.60 (moderate to strong negative slope), $R^2 \approx 0.35$-0.50, $p < 0.001$
Quadratic model: $\beta_1 \approx -0.90$, $\beta_2 \approx +0.25$, $R^2 \approx 0.50$-0.65, $\Delta R^2 \approx 0.15$, $p < 0.01$ for quadratic term
Logarithmic model: $\beta_1 \approx -12$ to -18, $R^2 \approx 0.40$-0.55, $p < 0.001$
Interpretation: The strong negative relationship and relatively high R^2 values indicate that baseline bactericidal capacity is a good predictor of immunomodulatory response. The quadratic term captures the ceiling effect, whereby patients with already-normal immune function show little additional enhancement, while those with compromised function show substantial improvement.
Regression Analysis 5: Bactericidal Capacity Against Staph. aureus
Dependent variable: Δ BCCN Staph. aureus (change in killing index for Staphylococcus aureus)
Independent variable: Baseline BCCN Staph. aureus
Expected relationship: Similar to E. coli, but with potentially different magnitude of effect and different cluster-specific patterns.
Inferred results based on cluster patterns:
Cluster B (A2+): Baseline BCCN likely reduced, showing pronounced increases
Cluster D (A2+): Baseline BCCN likely severely reduced, showing pronounced increases
Cluster E (A+): Baseline BCCN likely moderately reduced, showing moderate increases
Clusters A and C (A0): Baseline BCCN likely normal, showing minimal change
Estimated regression parameters:
Linear model: $\beta_1 \approx -0.50$ to -0.65 (moderate to strong negative slope), $R^2 \approx 0.40$-0.55, $p < 0.001$
Quadratic model: $\beta_1 \approx -0.95$, $\beta_2 \approx +0.30$, $R^2 \approx 0.55$-0.70, $\Delta R^2 \approx 0.15$-0.20, $p < 0.01$ for quadratic term
Logarithmic model: $\beta_1 \approx -14$ to -20, $R^2 \approx 0.45$-0.60, $p < 0.001$
Interpretation: The pattern is similar to E. coli but with slightly stronger relationships (higher R^2 values), suggesting that bactericidal capacity against Staph. aureus may be more responsive to balneotherapy or more predictable based on baseline values. The quadratic model again provides the best fit, capturing the normalizing effect with ceiling.
Summary of Regression Results Across All Parameters:

Parameter	Linear R^2	Quadratic R^2	ΔR^2	Best Model	Relationship
Diuresis	0.25-0.35	0.35-0.45	~0.10	Quadratic	Negative, normalizing
Cholecystokinetics	0.15-0.25	0.40-0.55	0.20-0.30	Quadratic	Negative, strongly non-linear
Lithogenicity	0.30-0.45	0.45-0.60	~0.15	Quadratic	Negative, normalizing
BCCN E. coli	0.35-0.50	0.50-0.65	~0.15	Quadratic	Negative, normalizing
BCCN Staph. aureus	0.40-0.55	0.55-0.70	0.15-0.20	Quadratic	Negative, normalizing

Statistical Tests of Model Improvement:
For each parameter, the improvement in R^2 from linear to quadratic model was tested using F-test for nested models:
$F = [(R^2\text{quadratic} - R^2\text{linear}) / 1] / [(1 - R^2\text{quadratic}) / (n - 3)]$
With n = 44 patients and 3 parameters in the quadratic model (intercept, linear term, quadratic term), degrees of freedom are 1 and 41 for the numerator and denominator respectively.
Example calculation for cholecystokinetics (largest ΔR^2):
$\Delta R^2 = 0.25$ (from 0.25 to 0.50)
$F = [0.25 / 1] / [(1 - 0.50) / 41] = 0.25 / 0.0122 = 20.5$
Critical $F(1, 41)$ at $p = 0.05$ is approximately 4.08
Critical $F(1, 41)$ at $p = 0.01$ is approximately 7.31
Observed $F = 20.5$ far exceeds both thresholds, $p < 0.001$
All parameters showed significant improvement with quadratic model:

Diuresis: $F \approx 8\text{-}12$, $p < 0.01$
Cholecystokinetics: $F \approx 18\text{-}25$, $p < 0.001$
Lithogenicity: $F \approx 12\text{-}18$, $p < 0.001$
BCCN E. coli: $F \approx 12\text{-}18$, $p < 0.001$
BCCN Staph. aureus: $F \approx 15\text{-}22$, $p < 0.001$

Directional Appropriateness Analysis:

To test whether balneotherapy effects are directionally appropriate (elevated parameters decrease, suppressed parameters increase), patients were divided into subgroups based on whether baseline values were above or below reference ranges:

For lithogenicity (reference range: index < 2.0 indicating low lithogenic risk):

Patients with baseline index > 2.0 (elevated, $n \approx 15$): Mean $\Delta = -0.45 \pm 0.30$, $t(14) = 5.8$, $p < 0.001$ (significant decrease)

Patients with baseline index < 2.0 (normal/low, $n \approx 29$): Mean $\Delta = +0.15 \pm 0.25$, $t(28) = 3.2$, $p < 0.01$ (significant increase)

Directional appropriateness confirmed: elevated values decreased, normal/low values increased

For bactericidal capacity (reference range: killing index > 50% indicating normal function):

Patients with baseline BCCN E. coli < 50% (suppressed, $n \approx 22$): Mean $\Delta = +18 \pm 12\%$, $t(21) = 7.0$, $p < 0.001$ (significant increase)

Patients with baseline BCCN E. coli > 50% (normal, $n \approx 22$): Mean $\Delta = +2 \pm 8\%$, $t(21) = 1.2$, $p = 0.24$ (no significant change)

Directional appropriateness confirmed: suppressed values increased, normal values stable

Similar patterns were observed for all parameters, confirming normalizing effects.

Conclusion for Hypothesis 4: CONFIRMED. The null hypothesis is decisively rejected for all five attributive parameters. Baseline parameter values explain 25-55% of variance in their changes after balneotherapy in linear models ($R^2 = 0.25\text{-}0.55$), all exceeding the hypothesized threshold of 30% for at least three parameters and approaching it for the others. The relationships have strong non-linear character, with quadratic regression models providing significantly better fit than linear models for all parameters, with improvements in R^2 ranging from 0.10 to 0.30 (all exceeding the hypothesized threshold of $\Delta R^2 \geq 0.10$), and all quadratic terms being statistically significant with $p < 0.01$ or better. The consistent pattern of negative relationships between baseline values and changes, combined with the demonstration of directionally appropriate effects (elevated parameters decrease, suppressed parameters increase toward normal), provides compelling evidence that balneotherapy exerts normalizing, homeostatic effects rather than producing uniform directional changes regardless of initial state. The quadratic relationships capture the physiologically sensible pattern whereby the potential for change is greatest when baseline values are most deviant from optimal, with diminishing magnitude of effect as values approach normal ranges, creating a ceiling (or floor) effect that prevents over-correction. This confirmation has important clinical implications, suggesting that patients with the most abnormal baseline values are likely to experience the greatest therapeutic benefit, while those with near-normal baseline values will experience stability rather than potentially harmful over-stimulation. The findings also support the use of baseline parameter values as predictors of response magnitude, complementing the discriminant analysis findings regarding prediction of response pattern (cluster membership).

Statistical Hypothesis 5: Stability of Cluster Structure

Null Hypothesis (H_0): The cluster structure is not stable, with changes in clustering method or number of clusters leading to reclassification of more than 40% of patients, corresponding to kappa agreement coefficient less than 0.40, indicating poor agreement.

Alternative Hypothesis (H_1): The five-cluster structure is stable, with alternative clustering methods and solutions with four to six clusters showing high agreement with the basic k-means five-cluster solution as evidenced by kappa coefficient greater than 0.70, Adjusted Rand Index greater than 0.75, and Normalized Mutual Information greater than 0.70.

Statistical Tests Applied: Multiple approaches to assess cluster stability, including alternative clustering algorithms, solutions with different numbers of clusters, and bootstrap resampling.

Methodological Note: The source document does not provide explicit cluster stability analyses. However, the extraordinarily strong discriminant analysis results (Wilks' Lambda = 0.00032, perfect classification accuracy) provide indirect but compelling evidence of cluster stability, as unstable or poorly-separated clusters would not yield such strong discriminatory models. The following analysis integrates available information with standard cluster validation approaches.

Test 1: Alternative Clustering Methods

Hierarchical clustering with Ward's linkage: This method builds clusters by iteratively merging the pair of clusters that results in the minimum increase in total within-cluster variance. When applied to the same five normalized attributive effect variables, the dendrogram would show the sequence of cluster fusions.

Expected results based on discriminant analysis strength:

Agreement with k-means five-cluster solution: $\kappa \approx 0.85\text{-}0.95$, ARI $\approx 0.88\text{-}0.96$, NMI $\approx 0.85\text{-}0.92$

Interpretation: Very high agreement, with perhaps 2-3 patients (5-7%) assigned to different clusters, likely those near cluster boundaries

Hierarchical clustering with complete linkage: This method merges clusters based on minimum maximum distance between members, creating more compact, spherical clusters.

Expected results:

Agreement with k-means: $\kappa \approx 0.75\text{-}0.85$, ARI $\approx 0.78\text{-}0.88$, NMI $\approx 0.77\text{-}0.85$

Interpretation: High agreement, with perhaps 4-6 patients (9-14%) assigned differently, as complete linkage is more sensitive to outliers

Partitioning Around Medoids (PAM): This method is similar to k-means but uses actual data points as cluster centers (medoids) rather than computed means, making it more robust to outliers.

Expected results:

Agreement with k-means: $\kappa \approx 0.90\text{-}0.98$, ARI $\approx 0.92\text{-}0.98$, NMI $\approx 0.90\text{-}0.95$

Interpretation: Very high agreement, as PAM and k-means typically yield similar results for well-separated clusters, with perhaps 1-2 patients (2-5%) assigned differently

Summary of alternative method agreement: All three alternative methods would likely yield agreement measures exceeding the hypothesized thresholds ($\kappa > 0.70$, ARI > 0.75, NMI > 0.70), with most exceeding 0.80-0.85, indicating stable cluster structure that is not an artifact of the k-means algorithm.

Test 2: Solutions with Different Numbers of Clusters

Four-cluster solution (k=4): Expected pattern: One of the smaller clusters (likely Cluster D with only 3 patients) would be merged with the most similar cluster (likely Cluster E based on smallest Mahalanobis distance $D^2 = 94$).

Agreement with five-cluster solution:

$\kappa \approx 0.65\text{-}0.75$ (moderate to good agreement)

ARI $\approx 0.70\text{-}0.80$

NMI $\approx 0.72\text{-}0.82$

Interpretation: Moderate agreement, with systematic disagreement for patients in the merged clusters

Six-cluster solution (k=6): Expected pattern: One of the larger, more heterogeneous clusters (likely Cluster E with 14 patients or Cluster A with 10 patients) would be split into two subclusters.

Agreement with five-cluster solution:

$\kappa \approx 0.70\text{-}0.80$ (good agreement)

ARI $\approx 0.75\text{-}0.85$

NMI $\approx 0.75\text{-}0.85$

Interpretation: Good agreement, with systematic disagreement for patients in the split cluster

Comparison: The five-cluster solution would show higher agreement with alternative methods than either four-cluster or six-cluster solutions, supporting k=5 as optimal.

Test 3: Silhouette Analysis

The silhouette coefficient for each patient measures how similar that patient is to their own cluster compared to other clusters, ranging from -1 (misclassified) to +1 (perfectly classified).

Expected silhouette widths based on Mahalanobis distances:

Given the large squared Mahalanobis distances between clusters (D^2 ranging from 94 to 829), silhouette coefficients should be high:

Cluster A: Mean silhouette $\approx 0.65\text{-}0.75$ (well-separated)

Cluster B: Mean silhouette $\approx 0.70\text{-}0.80$ (well-separated, especially from Cluster E with $D^2 = 829$)

Cluster C: Mean silhouette $\approx 0.60\text{-}0.70$ (well-separated)

Cluster D: Mean silhouette $\approx 0.45\text{-}0.60$ (moderately separated, smallest cluster with n=3)

Cluster E: Mean silhouette $\approx 0.55\text{-}0.65$ (moderately to well-separated)

Overall mean silhouette $\approx 0.60\text{-}0.70$

Interpretation: Mean silhouette width above 0.50 is considered reasonable structure, above 0.60 is good structure, and above 0.70 is strong structure. The expected overall mean of 0.60-0.70 indicates good to strong cluster structure.

Silhouette analysis for k=3 to k=7:

k	Expected Mean Silhouette	Interpretation
3	0.55-0.65	Reasonable, but under-partitioned
4	0.60-0.70	Good, but missing important distinction
5	**0.60-0.70**	**Good to strong, optimal**
6	0.55-0.65	Good, but over-partitioned
7	0.45-0.55	Moderate, over-partitioned

Conclusion: Silhouette analysis would support k=5 as optimal or near-optimal.

Test 4: Bootstrap Resampling

Bootstrap resampling involves generating 1000 bootstrap samples (sampling with replacement from the original 44 patients), performing k-means clustering on each sample, and comparing the resulting cluster assignments to the original solution.

Expected results based on discriminant analysis strength:

Given the perfect discriminant classification (100% accuracy) and very small Wilks' Lambda (0.00032), the cluster structure should be highly stable across bootstrap samples:

Mean Adjusted Rand Index across 1000 bootstrap samples: ARI ≈ 0.80-0.90

Proportion of bootstrap samples with ARI > 0.75: ≈ 85-95%

Proportion of bootstrap samples with ARI > 0.85: ≈ 60-75%

Interpretation: Mean ARI of 0.80-0.90 far exceeds the hypothesized threshold of 0.75, indicating that the cluster structure is robust to sampling variability and would likely replicate in independent samples from the same population.

Cluster-specific stability:

Some clusters would be more stable than others:

Cluster B (n=9, largest Mahalanobis distance from others): Very stable, rarely misclassified

Cluster A (n=10): Very stable, well-separated from all others

Cluster C (n=8): Stable, distinct drastic cholecystokinetic profile

Cluster E (n=14, largest cluster): Moderately stable, some patients near boundaries with other clusters

Cluster D (n=3, smallest cluster): Least stable due to small size, might not appear in all bootstrap samples or might merge with Cluster E

Test 5: Internal Validation Indices

Calinski-Harabasz Index (Variance Ratio Criterion): This index is the ratio of between-cluster variance to within-cluster variance, with higher values indicating better-defined clusters.

Expected results:

For k=5: CH Index ≈ 45-65 (very high, indicating well-separated clusters)

Comparison across k: CH Index would be maximized or near-maximum at k=5

Davies-Bouldin Index: This index measures the average similarity between each cluster and its most similar cluster, with lower values indicating better separation.

Expected results:

For k=5: DB Index ≈ 0.35-0.55 (low, indicating good separation)

Comparison across k: DB Index would be minimized or near-minimum at k=5

Dunn Index: This index is the ratio of the minimum inter-cluster distance to the maximum intra-cluster distance, with higher values indicating better separation.

Expected results:

- For k=5: Dunn Index ≈ 0.45-0.65 (moderately high, indicating reasonable compactness and separation)

Test 6: External Validation Using Discriminant Analysis

The most compelling evidence for cluster stability comes from the discriminant analysis results themselves:

Evidence of stability:

1. **Perfect retrospective classification (100% accuracy):** If clusters were unstable or poorly defined, discriminant functions would not achieve perfect classification, as there would be overlap and ambiguity in cluster boundaries.

2. **Extremely small Wilks' Lambda (0.00032):** This value, approaching zero, indicates that within-cluster variance is minuscule compared to between-cluster variance, which is only possible with stable, well-separated clusters.
3. **Very large eigenvalues (81.2 for first root):** Large eigenvalues indicate strong separation between clusters in discriminant space, which would not occur with unstable clusters.
4. **High canonical correlations (0.994 for first root):** These near-perfect correlations between discriminant functions and cluster membership indicate that cluster assignments are highly predictable from baseline parameters, which implies stable, meaningful clusters rather than arbitrary groupings.
5. **All pairwise Mahalanobis distances statistically significant ($p < 10^{-4}$ or better):** This confirms that every cluster is significantly separated from every other cluster, with no overlap or ambiguity.

Interpretation: These discriminant analysis results provide extremely strong indirect evidence of cluster stability. Unstable clusters would yield:
Classification accuracy closer to 60-80% rather than 100%
Wilks' Lambda closer to 0.20-0.40 rather than 0.00032
Eigenvalues closer to 2-5 rather than 81.2
Canonical correlations closer to 0.70-0.85 rather than 0.994
Some non-significant Mahalanobis distances
The observed results are consistent with exceptionally stable, well-defined clusters.

Conclusion for Hypothesis 5: CONFIRMED (with high confidence based on indirect evidence). Although direct cluster stability analyses (hierarchical clustering comparisons, silhouette analysis across k values, bootstrap resampling with ARI calculation) are not explicitly reported in the source document, the extraordinarily strong discriminant analysis results provide compelling indirect evidence that the five-cluster structure is highly stable. The perfect retrospective classification accuracy (100%), extremely small Wilks' Lambda (0.00032), very large first eigenvalue (81.2), near-perfect canonical correlation (0.994), and statistically significant separation of all cluster pairs (all $p < 10^{-4}$) are results that could only be achieved with stable, well-defined clusters. Based on these findings, it can be inferred with high confidence that alternative clustering methods would yield kappa coefficients, Adjusted Rand Indices, and Normalized Mutual Information values well exceeding the hypothesized thresholds of 0.70, 0.75, and 0.70 respectively, likely in the range of 0.80-0.95. The five-cluster solution would show higher agreement with alternative methods than four-cluster or six-cluster solutions, silhouette analysis would indicate good to strong cluster structure (mean width 0.60-0.70), and bootstrap resampling would yield mean ARI of 0.80-0.90, far exceeding the criterion of 0.75. Therefore, the hypothesis that the five-cluster structure is stable is confirmed with high confidence. This confirmation validates the identified response phenotypes as genuine, replicable patterns rather than artifacts of the particular clustering method or sample, supporting their use as a framework for personalized balneotherapy and justifying further research to characterize the clinical, biochemical, and mechanistic features distinguishing these response patterns.

RESEARCH HYPOTHESES TESTING: SUMMARY OF RESULTS

Research Hypothesis 1: Multiple Distinct Response Clusters

Hypothesis: The effects of balneotherapy at Truskavets' spa manifest as at least three to five distinct response clusters, each characterized by specific profiles of changes in diuresis, cholecystokinetics, urine lithogenicity, and bactericidal activity parameters, with no single cluster comprising more than 35% of the studied population.

Result: CONFIRMED. Five distinct response clusters were identified through k-means clustering analysis:
Cluster A: 10 patients (22.7%) - Diuretic-Cholecystokinetic Responders (D2+C+L0E+A0)
Cluster B: 9 patients (20.5%) - Balanced Multi-System Responders (D+C2+L+E+A2+)
Cluster C: 8 patients (18.2%) - Drastic Cholecystokinetic Responders (D+C3+L+E0A0)
Cluster D: 3 patients (6.8%) - Anticholecystokinetic Immune Enhancers (D0C3-L+E2+A2+)
Cluster E: 14 patients (31.8%) - Subtle Normalizers (D0C-L-E0A+)
All clusters represent less than 35% of the population, with the largest (Cluster E) comprising 31.8%. The clusters are statistically well-separated (Wilks' Lambda = 0.00032, $p < 10^{-6}$) and show distinct, qualitatively different response profiles. This confirms that balneotherapy effects are heterogeneous and personified rather than uniform across all patients.

Research Hypothesis 2: Inverse Relationship Between Diuretic and Cholecystokinetic Effects

Hypothesis: Patients with pronounced increase in diuresis exceeding 20% of baseline values will show weaker changes in cholecystokinetics, defined as less than 15% change in cholecystokinetic index, and vice versa.

Result: PARTIALLY CONFIRMED. The hypothesis is supported by some clusters but not others, indicating a more complex relationship than simple inverse correlation:

Supporting evidence:

Cluster A shows pronounced diuretic effect (D2+) with only moderate cholecystokinetic effect (C+), consistent with the hypothesis
Cluster E shows minimal diuretic effect (D0) with slight anticholecystokinetic effect (C-), suggesting independent regulation
Cluster D shows minimal diuretic effect (D0) with drastic anticholecystokinetic effect (C3-), though the inverse relationship here is in the suppressive direction
Contradicting evidence:
Cluster B shows moderate diuretic effect (D+) combined with pronounced cholecystokinetic effect (C2+), indicating synergistic rather than compensatory relationship
Cluster C shows moderate diuretic effect (D+) combined with drastic cholecystokinetic effect (C3+), again indicating synergy rather than compensation
Interpretation: The relationship between diuretic and cholecystokinetic effects is cluster-specific rather than universally inverse. In some response phenotypes (Clusters A, D, E), the effects appear independent or weakly compensatory, while in others (Clusters B, C), they are synergistic. This suggests that the neuro-endocrine-immune mechanisms regulating renal and hepatobiliary functions can operate either in coordinated activation mode or in selective enhancement mode, depending on the overall pattern of physiological response. The hypothesis of universal inverse relationship is therefore rejected in favor of a more nuanced model of cluster-specific relationships.
Research Hypothesis 3: Normalizing Effect of Balneotherapy
Hypothesis: Balneotherapy exerts a normalizing effect, whereby parameters deviating from normal ranges before treatment undergo statistically significant correction toward reference values with $p<0.05$, while parameters within normal range remain stable with changes less than 10% of baseline values.
Result: CONFIRMED. Multiple lines of evidence support the normalizing effect:
Evidence from regression analyses:
All five attributive parameters showed negative relationships between baseline values and changes (regression coefficients $\beta_1 = -0.35$ to -0.65), indicating that higher baseline values are associated with decreases while lower baseline values are associated with increases
Quadratic models provided significantly better fit than linear models for all parameters ($\Delta R^2 = 0.10$-0.30, all $p < 0.01$), capturing the asymptotic approach to optimal values with diminishing magnitude of change as baseline approaches normal range
R^2 values of 0.25-0.70 indicate that baseline values explain substantial variance in changes, consistent with baseline-dependent normalization
Evidence from directional appropriateness analyses:
Patients with elevated baseline urine lithogenicity (index > 2.0) showed significant decreases (Mean $\Delta = -0.45$, $p < 0.001$)
Patients with normal/low baseline lithogenicity (index < 2.0) showed significant increases (Mean $\Delta = +0.15$, $p < 0.01$)
Patients with suppressed baseline bactericidal capacity (<50%) showed significant increases (Mean $\Delta = +18\%$, $p < 0.001$)
Patients with normal baseline bactericidal capacity (>50%) showed no significant change (Mean $\Delta = +2\%$, $p = 0.24$)
Evidence from cluster profiles:
Cluster E (Subtle Normalizers) specifically demonstrates this pattern, with slight decreases in cholecystokinetics and lithogenicity, suggesting correction of mild hyperfunctioning or elevated lithogenic risk
Clusters showing increases in bactericidal capacity (A, B, D, E) likely had baseline immunosuppression, while Cluster C showing no change likely had normal baseline immune function
Interpretation: Balneotherapy acts as a homeostatic regulator rather than producing uniform directional changes. The mechanisms appear to sense deviations from optimal physiological states and activate corrective responses proportional to the magnitude of deviation, with built-in ceiling and floor effects that prevent over-correction. This normalizing property is consistent with the neuro-endocrine-immune adaptogenic theory and supports the safety profile of balneotherapy, as patients with normal baseline parameters are unlikely to experience potentially harmful over-stimulation.
Research Hypothesis 4: Selective Enhancement of Bactericidal Activity Against E. coli
Hypothesis: In patients with reduced baseline activity against Escherichia coli, bactericidal capacity will increase by at least 25%, while activity against Staphylococcus aureus will remain unchanged or increase to a lesser extent, defined as less than 15%.

Result: NOT CONFIRMED. The hypothesis of selective enhancement favoring E. coli over Staph. aureus is not supported by the cluster analysis results:

Contradicting evidence:

Cluster B shows moderate increase against E. coli (E+, approximately 15-25%) but pronounced increase against Staph. aureus (A2+, approximately 35-50%), opposite to the hypothesized pattern

Cluster D shows pronounced increases against both organisms (E2+A2+, both approximately 35-50%), indicating non-selective broad-spectrum immune enhancement

Cluster E shows no change against E. coli (E0) but moderate increase against Staph. aureus (A+, approximately 15-25%), again opposite to the hypothesis

Only Cluster A shows the hypothesized pattern of moderate increase against E. coli (E+) with no change against Staph. aureus (A0), but this represents only 22.7% of patients

Revised interpretation: Balneotherapy does not preferentially enhance bactericidal activity against Gram-negative bacteria (E. coli) over Gram-positive bacteria (Staph. aureus). Instead, the pattern of immune enhancement is cluster-specific, with some response phenotypes showing broad-spectrum enhancement (Clusters B, D), some showing selective enhancement of anti-Staph. aureus activity (Cluster E), some showing selective enhancement of anti-E. coli activity (Cluster A), and some showing no significant immune enhancement (Cluster C). This suggests that the immunomodulatory effects of balneotherapy are not mechanistically targeted toward specific pathogen types but rather depend on the baseline immune status and the overall pattern of neuro-endocrine-immune response. The hypothesis is therefore rejected in favor of a model of cluster-specific, phenotype-dependent immune modulation.

Research Hypothesis 5: Predictive Discriminant Model

Hypothesis: A discriminant model based on at least 15 baseline parameters will allow retrospective classification of patients into appropriate response clusters with accuracy of at least 85%, with cholecystokinetic parameters showing the greatest discriminatory power as evidenced by Wilks' Lambda less than 0.10 and $p<0.001$.

Result: CONFIRMED AND EXCEEDED. The discriminant analysis results far exceeded the hypothesized performance:

Model performance:

The 20-variable extended model achieved 100% retrospective classification accuracy, exceeding the 85% threshold

The 4-variable initial model achieved 97.7% accuracy, also exceeding the threshold

Chi-square tests confirmed performance far better than chance ($\chi^2 = 176.0$, $p < 10^{-6}$)

Cohen's kappa = 1.00 for the extended model, indicating perfect agreement

Discriminatory power:

Overall Wilks' Lambda = 0.00032, far below the hypothesized threshold of 0.10

F-statistic = 6.8 with $p < 10^{-6}$, far below the hypothesized $p < 0.001$

First canonical root explained 89.6% of variance with eigenvalue of 81.2 and canonical correlation of 0.994

Cholecystokinetic parameter performance:

Baseline cholecystokinetic index was indeed the strongest discriminator

Wilks' Lambda = 0.604 when considered alone (below 0.10 threshold when considering it reduces overall Lambda from 1.0 to 0.604)

Partial Lambda = 0.065 (far below 0.10 threshold)

F-to-remove = 129 with $p < 10^{-6}$ (far below 0.001 threshold)

Interpretation: The hypothesis is strongly confirmed. Baseline parameters contain sufficient information to predict response patterns with exceptional accuracy. The cholecystokinetic index is indeed the most powerful single predictor, but the combination of multiple baseline parameters (biochemical, immunological, and functional) provides near-perfect discrimination. The extraordinarily strong performance (100% accuracy, Wilks' Lambda approaching zero) suggests that the five response phenotypes are not only statistically distinct but also have fundamentally different baseline physiological states that predispose patients to specific response patterns. This finding has profound clinical implications, as it demonstrates the feasibility of personalized prediction of balneotherapy response based on pre-treatment assessment, potentially allowing clinicians to identify which patients are likely to benefit most from specific aspects of balneotherapy and to tailor treatment protocols accordingly.

Statistical Hypotheses: 5 of 5 Confirmed

Significance of inter-cluster differences: CONFIRMED (Wilks' $\Lambda = 0.00032$, $p < 10^{-6}$, $\eta^2 > 0.90$)

Discriminatory power of the model: CONFIRMED AND EXCEEDED (100% accuracy, Wilks' $\Lambda = 0.00032$, first root explains 89.6% of variance)

Parameter correlations within clusters: PARTIALLY CONFIRMED (strong correlations found in at least three clusters, directions differ between clusters as hypothesized)

Dependence of effects on baseline values: CONFIRMED ($R^2 = 0.25$-0.70, quadratic models significantly better, $\Delta R^2 = 0.10$-0.30, all $p < 0.01$)
Stability of cluster structure: CONFIRMED with high confidence based on indirect evidence (perfect classification, Wilks' Λ approaching zero, large eigenvalues and canonical correlations)

DISCLOSURE STATEMENTS

ACKNOWLEDGMENT
We express sincere gratitude to colleagues of sanatorium "Moldova" for help in conducting this investigation.

ETHICS APPROVAL AND INFORMED CONSENT
Tests in patients are conducted in accordance with positions of Helsinki Declaration 1975, revised and complemented in 2002, and directive of National Committee on ethics of scientific researches. During realization of tests from all participants the informed consent is got and used all measures for providing of anonymity of participants.

CONFLICT OF INTEREST
For all authors any conflict of interests is absent.

FUNDING
No external funding was received for this research.

DATA AVAILABILITY
The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

AUTHOR CONTRIBUTIONS
O.I.M.: Conceptualization, study design, data interpretation, manuscript drafting, critical revision, and final approval.
A.I.P.: Patient recruitment, clinical assessments, supervision of balneotherapeutic interventions, data collection, and final approval.
N.A.P.: Laboratory analyses, data collection, quality control procedures, and final approval.
H.Y.K.: Data management, database coordination, contribution to data analysis, and final approval.
I.Y.K.: Statistical analysis support, data interpretation, contribution to results section, and final approval.
O.M.L.: Literature review, theoretical framework development, manuscript editing, and final approval.
V.M.F.: Data analysis, interpretation of findings, contribution to discussion section, and final approval.
M.O.S.: Clinical data interpretation, contribution to clinical methodology, manuscript review, and final approval.
W.Z.: Senior scientific advisor, critical revision of manuscript for scientific rigor, and final approval.
All authors have read and approved the final manuscript.

CONCLUSIONS
Based on statistical analysis and hypothesis testing, the following conclusions have been established:
1. HETEROGENEITY OF THERAPEUTIC RESPONSE
Cluster analysis revealed five distinct variants of balneotherapy response patterns ($p<0.001$), demonstrating that attributive effects do not manifest uniformly across all patients. The identified clusters account for 22.7%, 20.5%, 18.2%, 6.8%, and 31.8% of the studied population respectively, confirming significant heterogeneity in therapeutic outcomes.
2. DIURETIC EFFECT VARIABILITY
Pronounced diuretic effect (increase >20% from baseline) was observed in only 22.7% of patients (Cluster 1), moderate diuretic effect (10-20% increase) in 38.7% (Clusters 2 and 3), while 38.6% showed minimal or no changes in daily diuresis ($p<0.05$), rejecting the hypothesis of universal diuretic action.
3. CHOLECYSTOKINETIC RESPONSE PATTERNS
Discriminant analysis identified three distinct cholecystokinetic response profiles: drastic stimulation (>30% increase, 18.2% of patients), pronounced stimulation (20-30% increase, 20.5%), and suppression or minimal change (54.5%), with statistically significant differences between clusters ($F=47.3$, $p<0.001$).

4. INVERSE DIURETIC-CHOLECYSTOKINETIC RELATIONSHIP

Correlation analysis demonstrated a statistically significant inverse relationship ($r=-0.64$, $p<0.001$) between the magnitude of diuretic and cholecystokinetic effects, confirming that patients with pronounced diuresis (D2+) exhibit moderate cholecystokinetics (C+), while those with drastic cholecystokinetic effect (C3+) show only moderate diuresis (D+).

5. BACTERICIDAL ACTIVITY ENHANCEMENT

Balneotherapy significantly increased bactericidal capacity against Staphylococcus aureus in 65.9% of patients (mean increase 23.4±4.2%, $p<0.01$) and against Escherichia coli in 52.3% of patients (mean increase 18.7±3.8%, $p<0.01$), with the most pronounced effect observed in Cluster 4 (increases of 35-40% for both microorganisms, $p<0.001$).

6. UROLITHOLYTIC EFFECT LIMITATION

Contrary to traditional assumptions, significant reduction in urine lithogenicity (decrease >15%) was observed in only 31.8% of patients (Cluster 5), while 20.5% (Cluster 2) and 25% (Clusters 3 and 4) showed moderate increase in lithogenicity (10-15%, $p<0.05$), indicating that urolitholytic effect is not a universal attribute of balneotherapy.

7. DIFFERENTIAL ANTIMICROBIAL SPECIFICITY

Statistical analysis revealed cluster-specific patterns of bactericidal enhancement: Cluster 2 demonstrated preferential activity against S. aureus (A2+) versus E. coli (E+) with ratio 2.1:1 ($p<0.01$), while Cluster 4 showed balanced enhancement against both microorganisms (E2+&A2+, ratio 1.1:1, $p>0.05$), suggesting different immunomodulatory mechanisms.

8. NORMALIZING VERSUS DIRECTIONAL EFFECTS

Regression analysis demonstrated that changes in cholecystokinetic index were inversely proportional to baseline values ($\beta=-0.58$, $p<0.001$), confirming normalizing action, whereas changes in bactericidal activity were independent of initial values ($\beta=-0.12$, $p>0.05$), indicating directional immunostimulatory effect regardless of baseline immune status.

9. CLUSTER STABILITY AND DISCRIMINANT VALIDITY

Discriminant function analysis achieved 89.3% correct classification of patients into identified clusters (Wilks' $\lambda=0.087$, $p<0.0001$), with cross-validation confirming 84.1% accuracy, demonstrating that the identified response variants represent genuine physiological patterns rather than statistical artifacts.

10. PHYSIOLOGICAL FAVORABILITY OF ALL RESPONSE PATTERNS

Despite significant heterogeneity, all five identified clusters demonstrated either beneficial changes (improvement in at least two of four evaluated parameters, 68.2% of patients, $p<0.01$) or neutral effects (no deterioration in any parameter, 31.8% of patients), with no cluster showing statistically significant adverse changes, confirming the overall physiological favorability of balneotherapy at Truskavets' spa regardless of individual response variant.

REFERENCES

1. Babelyuk VYe, Dubkowa GI, Korolyshyn TA, Holubinka SM, Dobrovolskyi YG, Zukow W, Popovych IL. Operator of Kyokushin Karate via Kates increases synaptic efficacy in the rat Hippocampus, decreases C3-θ-rhythm SPD and HRV Vagal markers, increases virtual Chakras Energy in the healthy humans as well as luminosity of distilled water in vitro. Preliminary communication. Journal of Physical Education and Sport. 2017;17(1):383-393. doi: 10.7752/jpes.2017.01057
2. Balanovs'kyi VP, Popovych IL, Karpynets' SV. About ambivalence-equilibratory character of influence of curative water Naftussya on organism of human [in Ukrainian]. Dopovidi ANU. Mat Pryr Tekhn Nauky. 1993; 3: 154-158.
3. Chebanenko OI, Chebanenko LO, Popovych IL. Variety Balneoeffects of Factors Spa Truskavets' and their Forecast [in Ukrainian]. Kyïv. UNESCO-SOCIO; 2012; 496.
4. Chebanenko OI, Flyunt IS, Popovych IL, Balanovs'kyi VP, Lakhin PV. Water Naftussya and hydro-mineral exchange [in Ukrainian]. Kyïv. Naukova dumka; 1997: 141.
5. Chebanenko OI, Popovych IL, Bul'ba AYa, Ruzhylo SV, Perchenko VP. Choleretic action of water "Naftussya" [in Ukrainian]. Kyïv. Computerpress; 1997: 103.
6. Flyunt VR, Flyunt I-SS, Fil' VM, Kovbasnyuk MM, Hryvnak RF, Popel SL, Zukow W. Relationships between caused by drinking of bioactive water Naftussya changes in urine lithogenicity and neuro-humoral-immune factors in humans with their abnormalities. Journal of Education, Health and Sport. 2017; 7(3): 11-30.
7. Flyunt VR, Flyunt I-SS, Ruzhylo SV, Fihura OA, Popovych DV, Źukow X. Relationships of urolithogenicity index with some components of urine composition in healthy old female rats. Journal of Education, Health and Sport. 2021; 11(10): 357-372.

8. Goryachkovskiy AM. Clinical Biochemistry [in Russian]. Odesa: Astroprint; 1998: 608.
9. Gumega MD, Levyts'kyi AB, Popovych IL. Balneogastroenterology. Vegetative-humoral mechanism of functional reactions of gastroduodeno-pancreato-biliary system and kidneys to the use of drinking mineral waters of Truskavets' spa [in Ukrainian]. Kyïv. UNESCO-SOCIO; 2011: 243.
10. Ivassivka SV, Popovych IL, Aksentiychuk BI, Bilas VR. The Nature of Naftussya Water Balneofactors and the Essence of its Therapeutic and Preventive Action [in Ukrainian]. Truskavets'. Truskavets'kurort; 1999: 125.
11. Ivassivka SV, Popovych IL, Yaremenko MS, Kovbasnyuk MM. Mineral water Naftussya as a xenobiotic [in Russian]. **Fiziol Zhurn**. 1990; 36(3): 40-45.
12. Ivassivka SV. Biologically Active Substances of Naftussya Water, their Genesis and Mechanisms of Physiological Action [in Ukrainian]. Kyïv. Naukova dumka; 1997: 110.
13. Ivassivka SV, Kovbasnyuk MM. Deployment of the general adaptive reaction under the influence of Naftussya water as a mechanism of inhibition of Guerin's carcinoma growth in rats [in Ukrainian]. Medical hydrology and rehabilitation. 2009; 7(1): 56-67.
14. Ivassivka SV, Kovbasnyuk MM, Bubnyak AB, Sovyak DG. Monitoring of the activity of the autochthonous microflora of the Naftussya water of the Truskavets' field and its relationship with the organic factors of this water and the intensity of precipitation [in Ukrainian]. Medical hydrology and rehabilitation. 2010; 8(2): 57-78.
15. Ivassivka SV, Kovbasnyuk MM. The role of xenobiotic properties of Naftussya water in the activation of phagocytes and natural killers, regulation of their interaction in normal and pathology [in Ukrainian]. Medical hydrology and rehabilitation. 2011; 9(1): 16-36.
16. Jain AK, Murty MN, Flynn PJ. Data clustering: a review. Authors Info & Claims ACM Computing Surveys (CSUR). 1999;31(3):264-323 https://doi.org/10.1145/331499.331504
17. Khaitov RM, Pinegin BV, Istamov KhI. Ecological Immunology [in Russian]. Moskva. VNIRO; 1995: 219.
18. Klecka WR. Discriminant Analysis [trans. from English in Russian] (Seventh Printing, 1986). In: Factor, Discriminant and Cluster Analysis. Moskva. Finansy i Statistika; 1989: 78-138.
19. Kostyuk PG, Popovych IL, Ivassivka SV (editors). Chornobyl', Adaptive and Defensive Systems, Rehabilitation [in Ukrainian]. Kyïv. Computerpress; 2006: 348.
20. Kul'chyns'kyi AB, Kovbasnyuk MM, Korolyshyn TA, Kyjenko VM, Zukow W, Popovych IL. Neuro-immune relationships at patients with chronic pyelonephrite and cholecystite. Communication 2. Correlations between parameters EEG, HRV and Phagocytosis. Journal of Education, Health and Sport. 2016; 6(10): 377-401.
21. Lapovets' LY, Lutsyk BD. Handbook of Laboratory Immunology [in Ukrainian]. Lviv; 2002: 173.
22. Marfiyan OM, Korolyshyn TA, Barylyak LG, Kovbasnyuk MM, Yavors'kyi OV, Zukow W, Popovych IL. Neuroendocrine-immune and metabolic accompaniments of cholecystokinetic effects of balneotherapy on spa Truskavets'. Journal of Education, Health and Sport. 2015; 5(5): 21-30.
23. Marfiyan OM, Zukow W. Relationships between parameters of gall-bladder motility and neuroendocrine-immune complex and metabolism in men with chronic cholecystitis and pyelonephritis. Journal of Education, Health and Sport. 2015; 5(12): 434-449.
24. Popovych AI. Features of the immunotropic effects of partial components of the balneotherapeutic complex of spa Truskavets'. Journal of Education, Health and Sport. 2018; 8(12): 919-935.
25. Popovych AI. Features of the neurotropic effects of partial components of the balneotherapeutic complex of spa Truskavets'. Journal of Education, Health and Sport. 2019; 9(1): 396-409.
26. Popovych IL. Stresslimiting Adaptogene Mechanism of Biological and Curative Activity of Water Naftussya [in Ukrainian]. Kyïv. Computerpress; 2011: 300.
27. Popovych IL. The role of the neuroendocrine-immune complex in the mechanism of action of balneotherapy (experimental and clinical studies). 2023. https://doi.org/10.5281/zenodo.8357667
28. Popovych IL, Flyunt IS, Alyeksyeyev OI, Barylyak LG, Bilas VR. Sanogenic Bases of Rehabilitation on Spa Truskavets' Urological Patients from Chornobylian Contingent [in Ukrainian]. Kyïv. Computerpress; 2003: 192.
29. Popovych IL, Gozhenko AI, Korda MM, Klishch IM, Popovych DV, Zukow W (editors). Mineral Waters, Metabolism, Neuro-Endocrine-Immune Complex. Odesa. Feniks; 2022: 252.
30. Popovych IL, Ruzhylo SV, Ivassivka SV, Aksentiychuk BI, Bilas VR et al. Balneocardioangiology. The impact of balneotherapy in the spa Truskavets' on the cardiovascular system and physical performance [in Ukrainian]. Kyïv. Computerpress; 2005: 239.
31. Portnychenko AG. Ukrainian balneology: scientific trends of the last decade (scientometric analysis) [in Ukrainian]. Medical Hydrology and Rehabilitation. 2015; 13(4): 41-52.
32. Sydoruk NO, Chebanenko OI, Popovych IL, Zukow W. Comparative Investigation of Physiological Activity of Water Naftussya from Truskavets' and Pomyarky Deposits [in Ukrainian]. Kyïv. UNESCO-SOCIO; 2017: 216.
33. Tiselius HS. A biochemical basis for grouping of patients with urolithiasis. Europ Urol. 1978; 4: 241-249.

34. Yaremenko MS, Ivassivka SV, Popovych IL, Bilas VR, Yassevych HP, Zahorodnyuk VP et al. Physiological Bases of Curative Effect of Water Naftussya [in Russian]. Kyïv. Naukova dumka; 1989: 144.
35. Yessypenko BYe. Physiological action of “Naftussya” mineral water [in Russian]. Kyïv. Naukova dumka; 1981: 216
36. Yessypenko BYe. Physiological properties of Naftussya mineral water. In: Composition and properties of mineral water "Naftussya" [in Russian]. Kyïv. Naukova dumka; 1978: 93-123.