Added value of a mental health specialist for evaluation of undiagnosed patients in centres for rare diseases – the ZSE-DUO cohort study

H. Hebestreit, A. Lapstich, C. Krauth, J. Deckert, K. Haas, L. Pfister, S. Witt, C. Schippers, J. Dieris-Hirche, T. Maisch, O. Tüscher, L. Bârlescu, A. Berger, M. Berneburg, V. Britz, A. Deibele, H. Graessner, H. Gündel, G. Heuft, Peter Heuschmann, T. Lücke, C. Mundlos, J. Quitmann, F. Rutsch, K. Schubert, J.B. Schulz, S. Schweiger, C. Zeidler, L. Zeltner, M. de Zwaan, and ZSE-DUO Working Group

Online Supplementary Material

Table of Contents

- Table S10) Proportion of patients with diagnoses fully explaining the symptomatic spectrum, including at least one non-rare somatic disease diagnosis.
- Table S11) Proportion of girls/women with diagnoses fully explaining the symptomatic spectrum.
- Table S12) Proportion of boys/men with diagnoses fully explaining the symptomatic spectrum.
- Table S13) Proportion of children/adolescents with diagnoses fully explaining the symptomatic spectrum.
- Table S14) Proportion of adults with diagnoses fully explaining the symptomatic spectrum.
- Table S15) Time to diagnoses fully explaining the symptomatic spectrum irrespective of the type of diagnoses (Mann-Whitney *U*-test).
- Table S16) Time to diagnoses fully explaining the symptomatic spectrum irrespective of the type of diagnoses (linear regression model).
- Table S17) Successful referral of patients with explaining diagnoses to local regular care in the standard care and innovative care cohorts relative to the respective entire cohort.
- Table S18) Successful referral of patients with explaining diagnoses to local regular care in the standard care and innovative care cohorts relative to the respective sample with at least one newly established diagnosis.

- Patients with at least one newly established diagnosis
  - Table S19) Proportion of patients with any newly established diagnosis.
  - Table S20) Proportion of patients with any newly established diagnosis, including at least one rare disease diagnosis.
  - Table S21) Proportion of patients with any newly established diagnosis, including at least one mental disorder diagnosis.
  - Table S22) Proportion of patients with any newly established diagnosis, including at least one non-rare somatic disease diagnosis.
  - Table S23) Time to first newly established diagnosis in patients with at least one newly established diagnosis (Mann-Whitney *U*-test).
  - Table S24) Time to first newly established diagnosis in patients with at least one newly establish diagnosis (linear regression model).
  - Table S25) Successful referral of patients with any new diagnosis to local regular care in the standard care and innovative care cohorts relative to the respective entire cohort.. (Fisher's exact test).
  - Table S26) Successful referral of patients with any new diagnosis to local regular care in the standard care and innovative care cohorts relative to the respective entire cohort (logistic regression model).
  - Table S27) Successful referral of patients with any new diagnosis to local regular care in the standard care and innovative care cohorts relative to the respective sample with at least one newly established diagnosis..
  - Table S28) Change in EQ-5D visual analogue scale rating between baseline and 12-month follow-up.
  - Table S29) Patient satisfaction with care at 12-month follow-up.

ZSE-DUO Working Group (collaborators)

Name	Surname
Federica	Akkaya
Christine	Babka
Lisa	Bannert
Anja	Bärsch-Michelmann
Leonie	Böhm
Folke	Brinkmann
Monika	Bullinger
Holger	Cario
Moritz	de Greck
Klaus-Michael	Debatin
Katrin	Dillmann-Jehn
Jutta	Eymann
Julia	Frisch
Anja	Glode
Vega	Gödecke
Corinna	Grasemann
Eva	Grauer
Astrid	Haas
Lea	Haisch
Isabell	Heinrich
Melissa	Held
Julia	Hennermann
Stephan	Herpertz
Anne	Herrmann-Werner
Julian	Hett
Peter	Heuschmann
Bettina	Hilbig
Laura	Holthöfer
Christiane	Imhof
Florian	Junne
Jan	Kassubek
Kevin-Thomas	Koschitzki
Heike	Krassort
Birgit	Kropff
Julia	Kuhn
Philipp	Latzko
Thomas	Loew
Albert C.	Ludolph
Torsten	Meyer
Isabell	Meyer dos Santos
Klaus	Mohnike

Martina	Monninger
Martin	Mücke
Susanne	Müller
Thomas	Musacchio
Margret	Nießen
Mariel	Nöhre
Stephan	Ott
Andrea	Petermann-Meyer
Christina	Pfeifer-Duck
Lea-Sophie	Piduhn
Carina	Rampp
Olaf	Rieß
Kristina	Schaubert
Annika	Schmidt
Simone	Schneider
Ludger	Schoels
Martina	Schwalba
Udo	Selig
Alexandra	Sroka
Toni	Steinbüchel
Sebastian	Stösser
Steffi	Suchant
Kathrin	Ungethüm
Matthias	Vogel
Daniela	Volk
Christoph	Vollmuth
Solange	Volnov
Thomas O. F.	Wagner
Sabrina	Walter
Bodo	Warrings
Kamil	Zajt
Karola	Zenker
David	Zhang
Stephan	Zipfel

Methods

ZSE-DUO is a prospective, controlled trial with a two-phase cohort design conducted in 11 centres of rare diseases (CRDs) in Germany (clinicaltrials.gov identifier: NCT03563677). Participating CRDs were located at the university hospitals in Aachen, Bochum, Frankfurt, Hannover, Magdeburg/Halle, Mainz, Münster, Regensburg, Tübingen, Ulm and Würzburg. All ethics committees of the participating CRDs and of the institutions involved in data analysis approved the project. Written informed consent was obtained from all participants and guardians, where applicable; all minors gave assent. Further description of the methodology is available in the published study protocol[1] and the original study protocol available at https://www.ukw.de/fileadmin/uk/zese/ZSE-DUO_Studienprotokoll_V1.3_04SEP2020.pdf.

Participants

Individuals aged 12 years or older who were referred by their treating physician for further diagnostic evaluation of a suspected rare disease to one of the participating CRDs were invited to participate in the trial. Treating physicians were required to provide a medical summary including reasons for suspecting a rare disease. Additional inclusion criteria were: 1) first contact with any of the participating CRDs, 2) attending the CRD's outpatient clinic for undiagnosed cases, and 3) providing written informed consent for study participation. Referrals were excluded from participation if 1) medical records available to the CRD were incomplete (i.e., missing medical summary letters, imaging results, blood tests etc.), or if 2) one or more disease(s) had previously been diagnosed, explaining the symptomatic spectrum presented. Furthermore, only patients insured by the statutory health insurance covering about 90% of the German population were included.

Standard care and innovative care

Standard care

The standard care (SC) cohort was recruited between October 2018 and September 2019. Once all required medical documents were available to the CRD, a summary document of the medical information was produced, and a multidisciplinary team discussed the case in a meeting. If the team concluded that no diagnoses covering the symptomatic spectrum were evident, the patient was invited to the CRD and seen in the outpatient clinic for undiagnosed cases by a physician with expertise in rare diseases and a specialisation in a 'somatic' medical discipline such as internal medicine, neurology, or paediatrics. As necessary, other disciplines could be involved and further assessments ordered (i.e., blood testing including genetic testing, imaging, neurography, etc.). Cases were discussed at least once more in a team meeting, and further diagnostic assessments and procedures could be requested. The referring physician and the patient received a letter summarising the findings and proposing future care and treatment.

Innovative care

The innovative care (IC) cohort, recruited between October 2019 and January 2021, received the same care as the SC group augmented by additional components. The major innovation was including a mental health specialist (physician with specialisation in psychiatry or psychosomatic medicine) in all aspects of the care process. The mental health specialist reviewed the patient's medical information upon admission to the CRD, evaluated the patient in the outpatient clinic in addition to the 'somatic' specialist and was – in close collaboration – involved in all decisions and actions taken. Mental disorders were diagnosed through extensive clinical evaluation using the diagnostic interview for mental disorders—the Mini-DIPS Open Access (OA).[2] The Mini-DIPS OA is a structured clinical interview for diagnosing mental disorders. Two training sessions were organised prior to the start of the IC to harmonise the diagnostic process between mental health experts across centres. We focused on current mental disorders. Depending on their findings, the mental health specialist could also offer up to ten face-to-face or teleconsultation sessions with the patient for further evaluation or to bridge the time to local mental health care. A supplementary component was nationwide case conferences among participating CRDs.

Study procedures

Data were collected from patients at three time points during the study: a) prior to the first visit to the outpatient clinic for undiagnosed cases (T0), b) on the day of the first visit to the CRD (T1), and c) 12 months after the first visit to the CRD (T2).

At baseline, we assessed sociodemographic data oriented to the survey instrument of the SOEP 2016,[3] signs and symptoms according to the human phenotype ontology (HPO),[4] and the past medical history, including all prior confirmed diagnoses. At baseline and 12-month follow-up, patients indicated their HRQoL on the visual analogue scale (VAS) of the EQ-5D-5L,[5] which ranges from 0 (worst imaginable health state) to 100 (best imaginable health state). EQ-5D-5L VAS reference data are available for the German adult population.[6] Additionally, at 12-month

follow-up, patients completed an established German questionnaire assessing satisfaction with care (ZUF-8).[7] The questionnaire comprises eight questions, each rated from 1 to 4 (total score 8 to 32).
Detailed information on study procedures is available in the published study protocol.[1]

Outcomes

Primary outcome

The primary outcome of the study was the proportion of patients with one or more diagnoses explaining their entire symptomatic spectrum. This outcome was assessed 12 months after the first visit to the CRD (T2). Physicians were asked to give their best judgment on whether all the pre-diagnoses and new diagnoses made during the intervention period fully explained the entire symptomatology (referred to as explaining diagnoses in this publication). In addition, they recorded each new diagnosis made during the intervention period and indicated whether it was a rare disease, a mental disorder, or a non-rare somatic disease.

Secondary outcomes

1) Time to diagnosis. Time to diagnosis was defined as the period between the initial visit to the CRD and the time the explaining diagnoses was made (in months). The date of first visit and diagnosis was sourced from the T2 physicians' questionnaire.
2) Transition to regular care. The success of the transition to regular care was defined as having attended a treatment appointment in a regular care setting following CRD recommendations. The number of patients successfully transferred was related to the total number of participants in the respective cohort. Recommendations regarding follow-up treatment were included in the data collected through the T2 physician's questionnaire. In contrast, the actual use of recommended treatment appointments was measured using data from both the patients' and the physicians' T2 surveys. In order to reduce the proportion of missing values, the details given by patients were supplemented with details given by the CRD physicians. A sensitivity analysis for the intersection of the information from the patient and doctor survey (N=116) resulted in an agreement rate of 86%. Data from 133 patients could be supplemented by n=197 data from the physicians.
3) Change in HRQoL from baseline to 12-month follow-up and satisfaction with care at 12-month follow-up.

Sample size calculation

We hypothesised that the IC would increase the proportion of patients with one or more diagnoses established during the work-up in the CRD, explaining the entire symptomatic spectrum of the patient from 30% in SC to 40% with IC. Power calculation was based on a Monte-Carlo simulation with 100.000 simulations for the assumed randomly varying centre-specific prevalence rates (between 20 and 40% with 95% probability on average) and odds-ratios (with 90% of CRDs with a positive effect on average). The calculated odds ratios were summarised using a random effects meta-analysis. Assuming a 20% dropout rate, a sample size of 682 patients in each group, 1,364 in total, was calculated to detect the above difference with a probability of a type 1 error <0.05 and a power of ≥ 0.8.

Statistical analysis

Highest school and post-secondary education were combined according to the International Standard Classification of Education (ISCED, 2011) for international comparability.[8] Open answers were recoded as far as possible into the existing categories or the categories summarised according to ISCED.
Reported net equivalent income was defined as net disposable household income (sum of earned, capital, transfer and other income of household members). A needs weighting was applied to household members by age, which was based on the modified OECD equivalence scale:[9] A weight of 1.0 was assigned to the head of household, a weight of 0.5 to each additional person aged 14 years or over, and a weight of 0.3 to each additional person younger than 14 years. Income was recorded in 21 groups. To generate an income distribution, a random value was generated for each person per group.
The net equivalent income was divided into three groups by the income thresholds 60 % and 200 % percent of the income median.[10] According to the current income median,[11], an income of less than € 1,251was determined for a low income, also called poverty risk, € 1,251to € 4,169 for a medium income and higher than € 4,169 for a high income.

Descriptive data are presented as numbers/proportions for nominal or ordinal variables and median/interquartile range (IQR) for continuous variables. For the primary outcome "explaining diagnoses", a mixed logistic regression model including a fixed study group effect along with random centre effects and random period effects nested within centres was employed. In a second step, the basic model was extended by adding demographic characteristics of patients (sex, age and highest post-secondary education) and in a third step by interaction terms between these characteristics and the study group. The two highest post-secondary education variables (Currently

enrolled in secondary/tertiary/vocational education and Bachelor's/postgraduate degree) were included after a significant chi-square test for differences between the SC and IC. The interaction effect for SC/IC*age was excluded due to multicollinearity with SC/IC. Data are provided as odds-ratios (OR) with 95% credibility intervals (95% CrI) for the main effects and 90% credibility intervals (90% CrI) for interactions. To reduce bias due to the skewed distribution of the dependent variable[12] and prevalence of 0 at centre level[13] in the sub-analyses by diagnosis type, the analyses were calculated using the Bayesian approach. The posterior distributions of the Bayesian model were estimated with a Markov chain Monte Carlo algorithm and uninformative priors by default of the analysis software (Stata 15.1).
According to Hox (2010),[12] effective sample size statistics (ESS) and the graphical output for autocorrelation plots as well as other measures for convergence diagnostics, were examined. For convergence diagnostics, the following graphical outputs were inspected according to the authors' recommendations: Tracing plots, Kernel density plots, cusum plots as well as the bivariate scatter plots of the model parameters based on the MCMC samples. In the basic model, these were satisfactory for an MCMC simulation rate of 10,000, burning rate of 2,500 and thinning rate of 50; in the other models with an MCMC simulation rate of 10,000, burning rate of 5,000 and thinning rate of 100. Despite the already very high simulation rate, the criterion of autocorrelation could not be achieved for four interaction models (subanalyses). For these models, a reduced model was calculated according to the recommendation of Hox (2010):[12] Non-significant interaction terms were removed or the model without interactions is to be assessed as the appropriate model.
Data are provided as OR with 95% CrI for the main effects and 90% credibility intervals (90% CrI) for interactions. The Bayes Factor (BF) and the Deviance Information Criterion (DIC) are reported to compare the multilevel models. The BF indicates the relative probability of how well a model fits the data compared to a (base) model. According to Kass & Raftery (1995),[14] a BF of 1 to 3 is considered a notable, 3 to 20 a positive, 20-150 a strong and greater than 150 a very strong improvement in model fit compared to the (base) model. A worse fitting model is indicated by a negative BF. The DIC is an index to indicate the deviation between hierarchical models. A smaller value indicates a more acceptable model.

Limitations of Mixed Models:
Nevertheless, biases cannot be completely ruled out. According to simulation studies, there is a risk of bias with a low cluster number and small cluster size, explicitly for a type I error for the fixed effect at level 2 and higher.[13,15] The authors of the studies recommend a sample size of at least 50 people at level one and a cluster size of 50 or 40 at level two.

For baseline data and secondary outcomes, differences between groups were tested using the $\chi 2$ test with Yates' continuity correction for 2 x 2 tables, Fisher's exact tests, or Mann–Whitney *U*-tests, according to the distribution of the variables. For significant differences, the effect sizes Cramér's *V* for proportion tests are reported, and the median differences in continuous outcomes between groups, including confidence intervals, were calculated using Hodges-Lehman approach are reported.
For the secondary outcome time to explaining diagnoses and HRQoL, a Mann–Whitney *U* -test was calculated, as the prerequisites of normal distribution and variance homogeneity of a t-test were not fulfilled.
For the secondary outcomes, regression models were also examined to control for the named adjustment variables. Linear multiple regression models were calculated for time to explaining diagnoses. For this purpose, the prerequisites of linearity of the examined correlations, multicollinearity, homoskedasticity and normal distribution of the residuals were tested. The conditions of homoscedasticity and normal distribution of the residuals could not be confirmed. To reduce bias in the point estimates and standard errors, bootstrapping with 5,000 draws was used in the estimation of the regression models.[12] To further correct for skewed distribution, bias-corrected and accelerated confidence intervals are reported on the effects. Outliers were retained due to the limited observation period of 12 months. Regression coefficients are reported to assess the effect. For model goodness, the Wald χ^2 model test, the adjusted R^2 is reported. A good model fit for a significant Wald χ^2 is found. R^2 is interpreted according to Cohen (1988)[16] for a value of 0.02 as weak, 0.13 as medium and 0.26 as high variance explanation.
For a model comparison, a Liklihood Ratio χ^2 as well as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are reported. A significant Liklihood Ratio χ^2 indicates a difference between the two models under comparison. AIC and BIC are used to interpret whether the model is worsening or improving. AIC and BIC are to be interpreted similarly to DIC: A smaller value here also indicates a more acceptable model.
For proportion of patients who were referred to (local) regular care multiple logistic regression models were calculated for the adjustment variables. The effects are reported as OR. A Wald χ^2 test, Nagelker's pseudo-R^2, Hosmer-Lemeshow χ^2 test and area under the curve (AUC) of a receiver operating characteristic (ROC) analysis are reported for the model fit. According to the distribution, a classification cutoff of 0.8 was chosen. A good model fit exists when Wald's χ^2 test becomes significant, the Hosmer-Lemeshow χ^2 test is not significant and the AUC value is > 0.7. According to Backhaus et al. (2021), Nagelkerger's pseudo-R^2 is rated as acceptable for an R^2 > 0.2, > 0.4 as good and > 0.5 as very good.

We performed linear multiple regression models for analysing changes in HRQoL from baseline to 12-month follow-up and patient satisfaction at 12-month follow-up. Analogous to the regression analyses for time to explaining diagnosis, we tested the prerequisites of linearity of the examined correlations, multicollinearity, homoskedasticity and normal distribution of the residuals. The analysis did not confirm the conditions of homoscedasticity and normal distribution of the residuals. Therefore, we used bootstrapping with 5,000 draws in estimating the regression models,[12] including bias-corrected and accelerated (Bca) confidence intervals. Regression coefficients are provided to evaluate the effect. The adjusted R^2 value is given for model goodness. We interpreted R^2 according to Cohen (1988). AIC and BIC are reported for the models and interpreted as mentioned above.

Participants with missing data were excluded from the respective analyses. Statistical significance was assumed at $p<0.05$ (two-sided test).
The statistical analyses were done in SAS, Stata 15.1 and SPSS 27 & 28.

References

1. Hebestreit H, Zeidler C, Schippers C, et al. Dual guidance structure for evaluation of patients with unclear diagnosis in centers for rare diseases (ZSE-DUO): study protocol for a controlled multi-center cohort study. *Orphanet J Rare Dis* 2022; **17**: 47. doi: 10.1186/s13023-022-02176-1.
2. Margraf J, Cwik JC. Mini-DIPS open access: Diagnostisches Kurzinterview bei psychischen Störungen. [Mini-DIPS open access: diagnostic interview for mental disorders]. Bochum: Forschungs-und Behandlungszentrum für psychische Gesundheit, Ruhr-Universität Bochum, 2017.
3. DIW Berlin / SOEP. (2017). SOEP-IS 2016: Fragebogen für die SOEP-Innovations-Stichprobe: SOEP Survey Papers No. 515. In *DIW Berlin*. Available at: https://www.diw.de/documents/publikationen/73/diw_01.c.789421.de/diw_ssp0865.pdf. Retrieved 1 March 2018.
4. Köhler S, Gargano M, Matentzoglu N, et al. The human phenotype ontology in 2021. *Nucleic Acids Res* 2021; **49**: D1207–17. doi: 10.1093/nar/gkaa1043.
5. Feng YS, Kohlmann T, Janssen MF, Buchholz I. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. *Qual Life Res* 2021; ***30***: 647–73. doi: 10.1007/s11136-020-02688-y.
6. Grochtdreis T, Dams J, König HH, Konnopka A. Health-related quality of life measured with the EQ-5D-5L: estimation of normative index values based on a representative German population sample and value set. *Eur J Health Econ* 2019; **20**: 933–44. doi: 10.1007/s10198-019-01054-1.
7. Schmidt J, Lamprecht F, Wittmann WW. Zufriedenheit mit der stationären Versorgung. Entwicklung eines Fragebogens und erste Validitätsuntersuchungen [Satisfaction with inpatient management. Development of a questionnaire and initial validity studies]. *Psychother Psychosom Med Psychol* 1989; **39**: 248–55.
8. UNESCO Institute for Statistics. International Standard Classification of Education - ISCED 2011. Montreal/Quebec, Canada 2012. Available at: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf. Retrieved 25 March 2023.
9. OECD. What are equivalence scales? Available at: https://www.oecd.org/economy/growth/OECD-Note-EquivalenceScales.pdf. Retrieved 18 March 2023.
10. Bundesministerium für Arbeit und Soziales (2022): Der Sechste Armuts- und Reichtumsbericht der Bundesregierung, Berlin. Available at: https://www.armuts-und-reichtumsbericht.de/SharedDocs/Downloads/Berichte/sechster-armuts-reichtumsbericht.pdf?__blob=publicationFile&v=6. Retrieved 12 March 2023.
11. Statistisches Bundesamt (Destatis) (2022): Lebensbedingungen und Armutsgefährdung: Einkommensverteilung (Nettoäquivalenzeinkommen). Available at:. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Lebensbedingungen-Armutsgefaehrdung/Tabellen/einkommensverteilung-mz-silc.html. Retrieved 12 March 2023.
12. Hox JJ. Multilevel Analysis: Techniques and Applications (2nd ed.). 2010. New York: Routledge. https://doi.org/10.4324/9780203852279.
13. Moineddin R, Matheson FI, Glazier RH. A simulation study of sample size for multilevel logistic regression models. *BMC Med Res Methodol* 2007; **7**: 34. doi: 10.1186/1471-2288-7-34.
14. Kass RE, Raftery AE. Bayes factors. J Am Stat Assos 1995; **90**: 773–795.
15. Schoeneberger JA. The impact of sample size and other factors when estimating multilevel logistic models. *J Exp Educ* 2016; **84**, 373–397. DOI: doi: 10.1080/00220973.2015.1027805

16. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Routledge. 1988. https://doi.org/10.4324/9780203771587
17. Backhaus K, Erichson B, Gensler S, Weiber R, Weiber T. *Multivariate Analysis. An Application-Oriented Introduction*. 1st Ed. Wiesbaden, Gernay: Springer Gabler. 2021. https://doi.org/10.1007/978-3-658-32589-3

Supplementary Figures

Patients with newly established diagnoses explaining the entire symptomatic spectrum presented

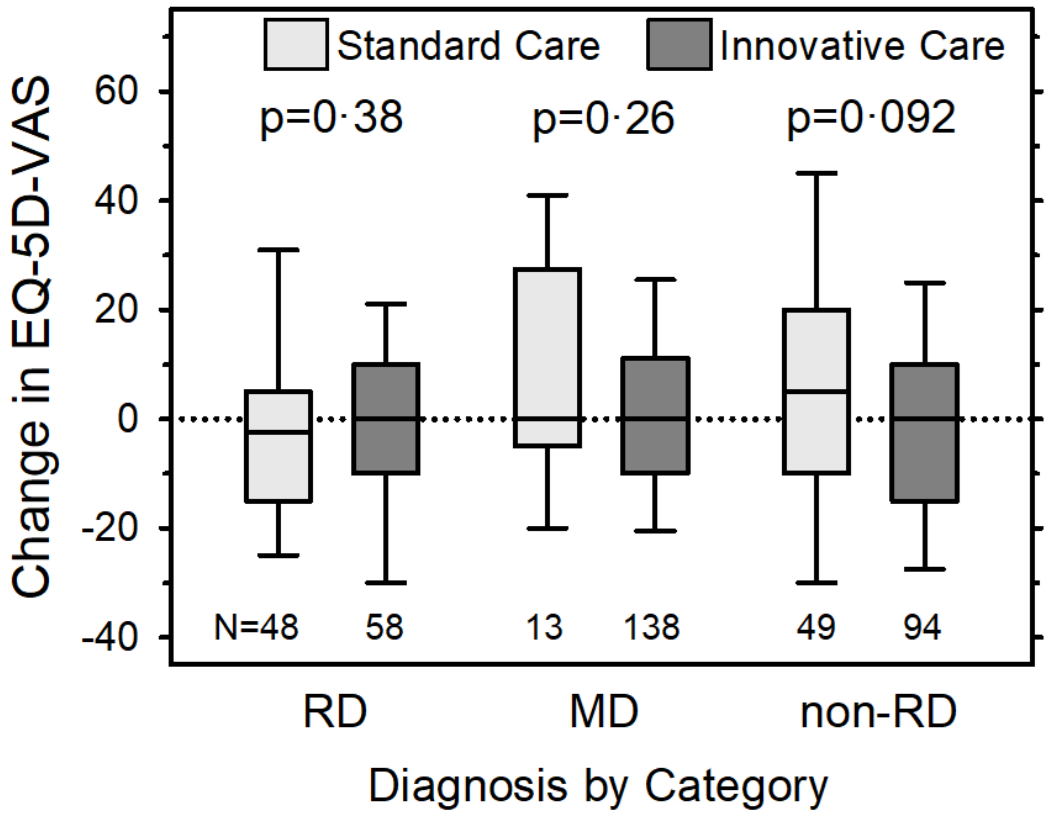

Figure S1) Change in health-related quality of life as indicated on the EQ-5D visual analogue scale between baseline and 12-month follow-up in patients with explaining diagnoses of the standard care and the innovative care cohorts.
Patients are clustered by diagnostic category. Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.
RD – rare disease; MD – mental disorder; non-RD – non-rare somatic disease.
Boxplots show 10th, 25th, 50th, 75th, and 90th centiles. Two-sided Mann-Whitney *U*-test was used for statistical analyses.

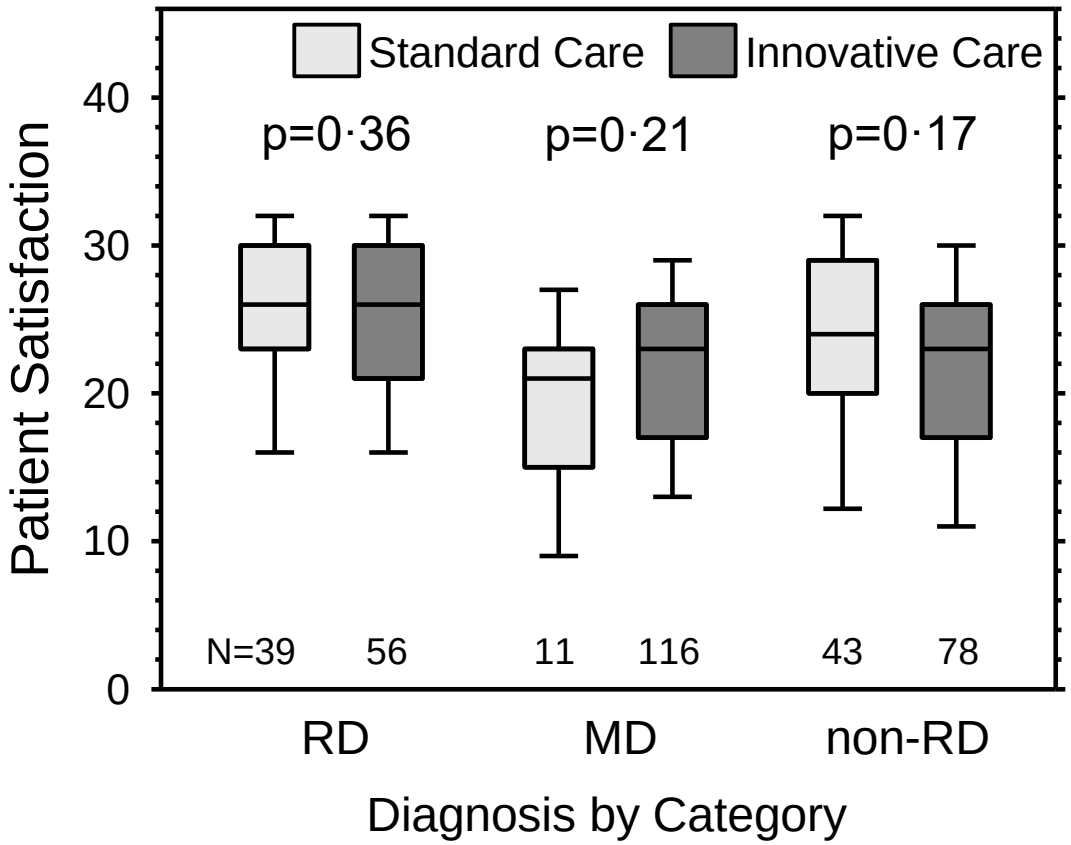

Figure S2) Patient satisfaction with care at 12-month follow-up in patients of the standard care and the innovative care cohorts with explaining diagnoses.
Patients are clustered by diagnostic category. Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.
RD – rare disease, MD – mental disorder, non-RD – non-rare somatic disease.
Boxplots show 10th, 25th, 50th, 75th, and 90th centiles. Two-sided Mann-Whitney *U*-test was used for statistical analyses.

Patients with at least one newly established diagnosis

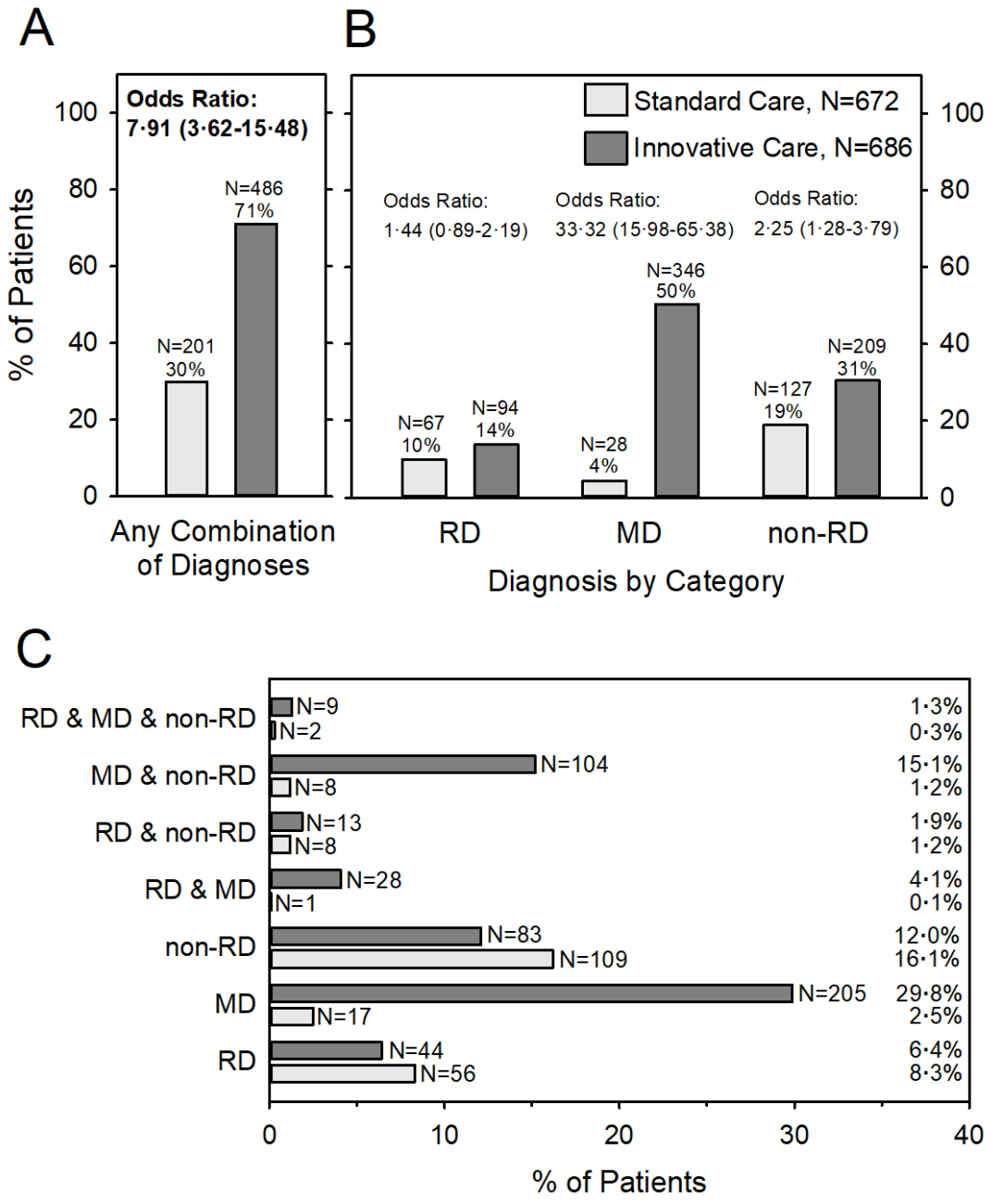

Figure S3) Proportion of patients in the standard care and the innovative care groups for whom at least one new diagnosis could be established during the evaluation process.

A) Patients with any newly established diagnoses.
B) Patients clustered by diagnostic category. Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.
C) Combinations of diagnostic categories.

RD – rare disease; MD – mental disorder; non-RD – non-rare somatic disease.
Main effects are presented as odds ratios and 95%-credibility intervals based on basic statistical models.

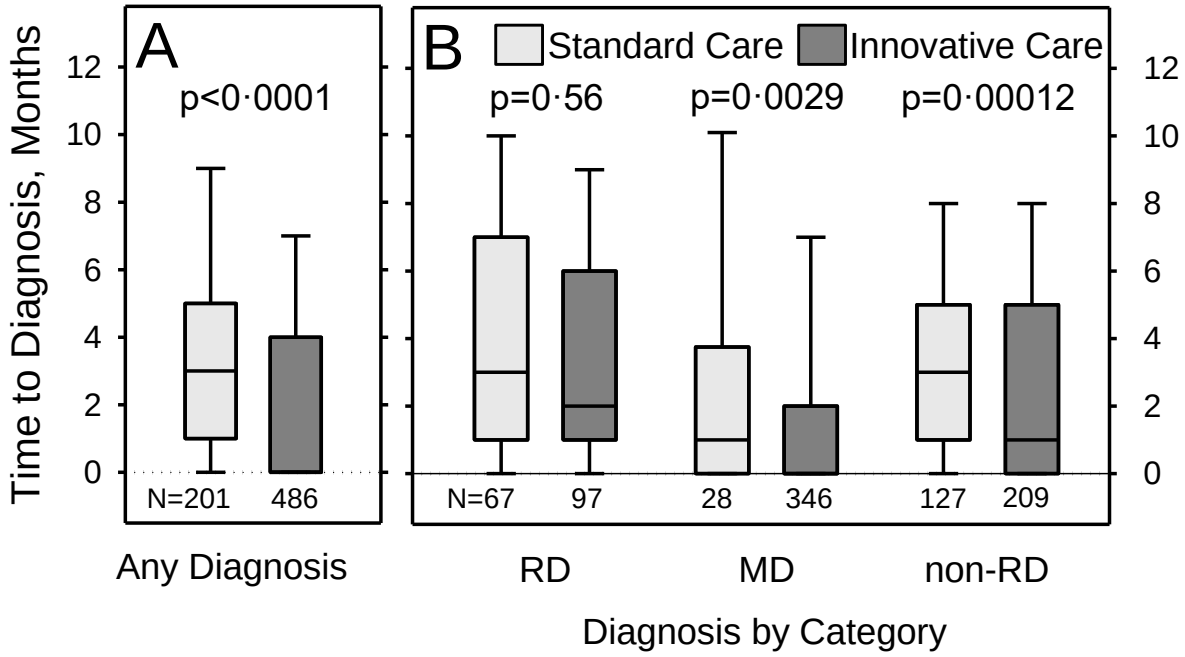

Figure S4) Time between first visit to the CRD and any newly established diagnosis in the standard care and innovative care cohorts.

A) Patients with any newly established diagnoses.
B) Patients clustered by diagnostic category. Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

RD – rare disease; MD – mental disorder; non-RD – non-rare somatic disease.
Boxplots show 10th, 25th, 50th, 75th, and 90th centiles. Two-sided Mann-Whitney *U*-test was used for statistical analyses.

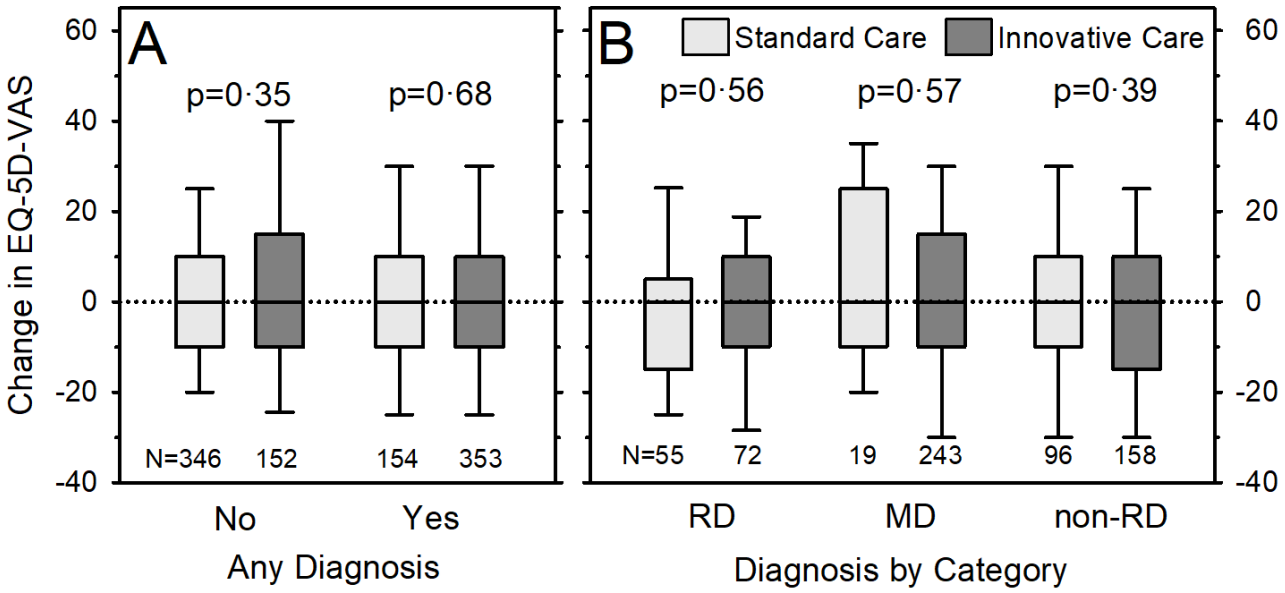

Figure S5) Change in EQ-5D-VAS between baseline and 12-month follow-up in the standard care and the innovative care cohorts.

A) Patients without and with any newly established diagnosis.
B) Patients with any newly established diagnosis clustered by diagnostic category. Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

RD – rare disease; MD – mental disorder; non-RD – non-rare somatic disease.
Boxplots show 10th, 25th, 50th, 75th, and 90th centiles. Two-sided Mann-Whitney *U*-test was used for statistical analyses.

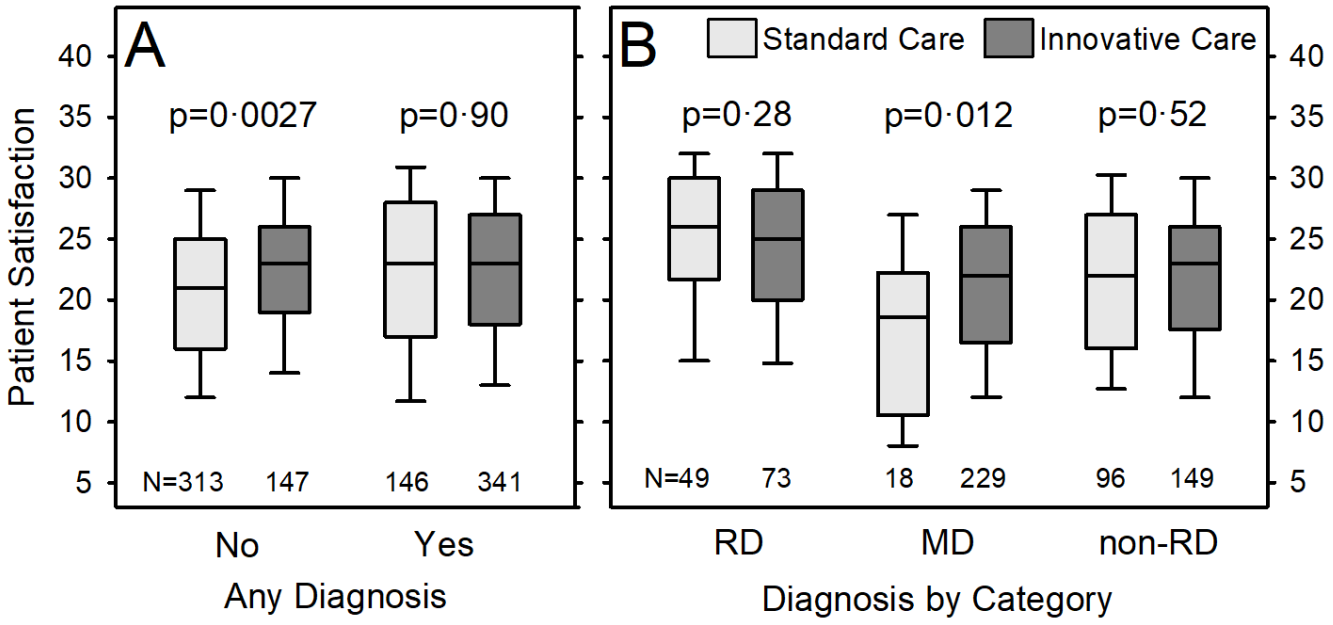

Figure S6) Patient satisfaction with care at 12-month follow-up in the standard care and the innovative care cohorts.

A) Patients without and with any newly established diagnosis
B) Patients with any newly established diagnosis clustered by diagnostic category. Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

RD – rare disease; MD – mental disorder; non-RD – non-rare somatic disease.
Boxplots show 10th, 25th, 50th, 75th, and 90th centiles. Two-sided Mann-Whitney *U*-test was used for statistical analyses.

Supplementary Tables

Descriptive information at baseline

Table S1) Patient characteristics at baseline comparing participants of the study to non-participants who were eligible but declined to participate.

	Participants N=1375	non participation N=182	*p* value[a]
Age (years) [median, IQR]	47 (33-58)	49 (35-59)	0·26
Disease duration (years) [median, IQR]	5 (2-12)	4 (2-10)	0·058
female sex [n, %]	832 (60·5)	122 (67·0)	0·11

[a] Chi-square test with Yates' continuity correction for 2 × 2 tables (sex), Non-parametric test (Mann-Whitney *U*-test) for numerical variables,

Table S2) Sociodemographic characteristics of participants at baseline.

	Standard care cohort (N=678)	Innovative care cohort (N=689)	*p* value[a]
Seen by a mental health specialist in the last 12 months [n,%]	236 (34·8)	264 (38·3)	0.23
Missing [n,%]	9 (1·3)	5 (0·7)	
Born outside Germany [n, %]	89 (13·1%)	73 (10·6%)	0·16
Declined to answer [n, %]	2 (0·3%)	0 (0·0%)	
Missing [n, %]	7 (1·0%)	4 (0·6%)	
Area of residence			0·12
Rural [n, %]	334 (49·3%)	371 (53·8%)	
Urban [n, %]	332 (49·0%)	309 ((44·8%)	
Declined to answer [n, %]	1 (0·1%)	0 (0·0%)	
Missing [n, %]	11 (1·6%)	9 (1·3%)	
Assistance in everyday life			1·00
Do not receive assistance [n, %]	575 (84·8%)	591 (85·8%)	
Receive assistance [n, %]	85 (12·5%)	87 (12·6%)	
Declined to answer [n, %]	11 (1·6%)	7 (1·0%)	
Missing [n, %]	7 (1·0%)	4 (0·6%)	
Equivalised disposable income[b]			0·52
At risk of poverty: < 1,251 €/month [n, %]	317 (46·8%)	301 (43·7%)	
Not at risk of poverty: >=1,251 - 4,169 €/month [n, %]	238 (35·1%)	246 (35·7%)	
Declined to answer [n, %]	108 (15·9%)	125 (18·1%)	
Missing [n, %]	15 (2·2%)	17 (2·5%)	

[a] Chi-Square test for categorical variables or chi-square test with Yates' continuity correction for 2 × 2 tables. Without the categories "Declined to answer" and "Missing".

[b] the mean net equivalent income in Germany in 2021 was 2,424 Euro/month, the respective median was 2,085 Euro/month (Statistisches Bundesamt 2022). The cut-off for the risk of poverty is defined as 60% of the median (€ 2,084·58 * 0·6 = € 1,250·75; Bundesministerium für Arbeit und Soziales 2022, p. 525).

References

Statistisches Bundesamt (Destatis) (2022): Lebensbedingungen und Armutsgefährdung: Einkommensverteilung (Nettoäquivalenzeinkommen). Available at: https://www·destatis·de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Lebensbedingungen-Armutsgefaehrdung/Tabellen/einkommensverteilung-mz-silc·html. Retrieved 12 March 2023.

Bundesministerium für Arbeit und Soziales (2022): Der Sechste Armuts- und Reichtumsbericht der Bundesregierung, Berlin. Available at: https://www·armuts-und-reichtumsbericht·de/SharedDocs/Downloads/Berichte/sechster-armuts-reichtumsbericht·pdf?__blob=publicationFile&v=6. Retrieved 12 March 2023.

Table S3) Participants' characteristics at baseline in female and male participants

	Female			Male		
	Standard care cohort (N=404)	Innovative care cohort (N=421)	*p* value[a]	Standard care cohort (N=274)	Innovative care cohort (N=268)	*p* value[a]
Age (years) [median, IQR]	47 (33-56)	46 (32-57)	0·89	46 (32-58)	47 (33-57)	0·94
Children/adolescents (≥12 to <18 years) [n, %]	19 (4·7%)	14 (3·3%)	0·41	23 (8·4%)	11 (4·1%)	0·060
Migratory background			0·71			0·22
No migratory background [n, %]	291 (72·0%)	301 (71·5%)		183 (66·8%)	194 (72·4%)	
Migratory background [n, %]	101 (25·0%)	111 (26·4%)		84 (30·7%)	69 (25·7%)	
Declined to answer [n, %]	8 (2·0%)	6 (1·4%)		4 (1·5%)	4 (1·5%)	
Missing [n, %]	4 (1·0%)	3 (0·7%)		3 (1·1%)	1 (0·4%)	
Highest school education			0·72			0·14
No graduation [n, %]	23 (5·7%)	19 (4·5%)		32 (11·7%)	19 (7·1%)	
Lower secondary (ISCED 2)[b] [n, %]	222 (55·0%)	236 (56·1%)		140 (51·1%)	128 (47·8%)	
Upper secondary (ISCED 3)[b] [n, %]	146 (36·1%)	157 (37·3%)		100 (36·5%)	110 (41·0%)	
Other educational degree [n, %]	2 (0·5%)	0 (0·0%)		1 (0·4%)	5 (1·9%)	
Missing [n, %]	11 (2·7%)	9 (2·1%)		1 (0·4%)	6 (2·2%)	
Highest post-secondary education			0·17			0·16
Currently enrolled in secondary/ tertiary/vocational education [n, %]	52 (12·9%)	40 (9·5%)		40 (14·6%)	27 (10·1%)	
No tertiary/vocational education, not currently enrolled [n, %]	31 (7·7%)	29 (6·9%)		17 (6·2%)	21 (7·8%)	
Vocational qualification (ISCED 4)[b] [n, %]	209 (51·7%)	209 (49·6%)		120 (43·8%)	106 (39·6%)	
Bachelor's/postgraduate degree (ISCED 5-8)[b] [n, %]	100 (24·8%)	129 (30·6%)		88 (32·1%)	105 (39·2%)	
Other tertiary/vocational degree [n, %]	0 (0·0%)	1 (0·2%)		1 (0·4%)	2 (0·7%)	
Declined to answer [n, %]	12 (3·0%)	13 (3·1%)		8 (2·9%)	7 (2·6%)	
Missing [n, %]	0 (0·0%)	0 (0·0%)		0 (0·0%)	0 (0·0%)	
Employment status			0·18			0·41
Full-time [n, %]	95 (23·5%)	111 (26·4%)		110 (40·1%)	120 (44·8%)	
Part-time or less [n, %]	74 (18·3%)	92 (21·9%)		15 (5·5%)	17 (6·3%)	
Unemployed [n, %]	40 (9·9%)	35 (8·3%)		33 (12·0%)	25 (9·3%)	
Retired due to disability [n, %]	64 (15·8%)	46 (10·9%)		27 (9·9%)	21 (7·8%)	
Outside of the labour force for other reasons [n, %]	100 (24·8%)	109 (25·9%)		71 (25·9%)	54 (20·1%)	
Declined to answer [n, %]	31 (7·7%)	28 (6·7%)		18 (6·6%)	31 (11·6%)	
Missing [n, %]	0 (0·0%)	0 (0·0%)		0 (0·0%)	0 (0·0%)	
Duration of main symptom (years) [median, IQR]	6 (3-15)	7 (3-16)	0·75	6 (3-14)	5 (2-14)	0·20
Missing [n, %]	2 (0·5%)	7 (1·7%)		0 (0·0%)	0 (0·0%)	
Number of HPO codes per patient [median, IQR]	6 (3-9)	7 (3-16)	0·66	6 (3-9)	5 (3-8)	0·77
Missing [n, %]	4 (1·0%)	4 (0·9%)		2 (0·7%)	0 (0·0%)	
Disability formally acknowledged [n, %][c]	153 (37·9%)	162 (38·5%)	0·95	96 (35·0%)	93 (34·7%)	0·91
Declined to answer [n, %]	10 (2·5%)	10 (2·4%)		9 (3·3%)	7 (2·6%)	
HRQoL: EQ-5D-VAS [median, IQR]	50 (30-70)	50 (30-65)	0·43	50 (30-70)	50 (35-70)	0·42
Missing [n, %]	0 (0·0%)	0 (0·0%)		0 (0·0%)	1 (0·4%)	

Abbreviations: HPO = human phenotype ontology, HRQoL = health-related quality of life, VAS = visual analogue scale (0-100)

[a] Chi-square test for categorical variables or chi-square test with Yates' continuity correction for 2 × 2 tables. Without the categories "Declined to answer", "Missing" and the open categories "Other tertiary/vocational degree" and "Other educational degree". Non-parametric test (Mann-Whitney *U*-test) for numerical variables,

[b] ISCED – International Standard Classification of Education 2011,

[c] reported by patients

Table S3) Participants' characteristics at baseline in female and male participants continued

	Female			Male		
	Standard care cohort (N=404)	**Innovative care cohort (N=421)**	***p* value[a]**	**Standard care cohort (N=274)**	**Innovative care cohort (N=268)**	***p* value[a]**
Seen by a mental health specialist in the last 12 months [n,%]	138 (34·2%)	157 (37·3%)	0·48	98 (35·8%)	107 (39·9%)	0·33
Missing [n, %]	8 (2·0%)	2 (0·5%)		1 (0·3%)	3 (1·1%)	
Born outside Germany [n, %]	60 (14·9%)	44 (10·5%)	0·065	29 (10·6%)	29 (10·8%)	1·00
Declined to answer [n, %]	4 (1·0%)	3 (0·7%)		3 (1·1%)	1 (0·4%)	
Missing [n, %]	2 (0·5%)	0 (0·0%)		0 (0·0%)	0 (0·0%)	
Area of residence [n, %]			0·69			0·055
Rural [n, %]	204 (50·5%)	219 (52·0%)		130 (47·4%)	152 (56·7%)	
Urban [n, %]	195 (48·3%)	196 (46·6%)		137 (50·0%)	113 (42·2%)	
Declined to answer [n, %]	1 (0·2%)	0 (0·0%)		0 (0·0%)	0 (0·0%)	
Missing [n, %]	4 (1·0%)	6 (1·4%)		7 (2·6%)	3 (1·1%)	
Assistance in everyday life [n, %]			0·90			0·94
Do not receive assistance [n, %]	335 (82·9%)	357 (84·8%)		240 (87·6%)	234 (87·3%)	
Receive assistance [n, %]	56 (13·9%)	57 (13·5%)		29 (10·6%)	30 (11·2%)	
Declined to answer [n, %]	9 (2·2%)	4 (1·0%)		2 (0·7%)	3 (1·1%)	
Missing [n, %]	4 (1·0%)	3 (0·7%)		3 (1·1%)	1 (0·4%)	
Equivalised disposable income[d] [n, %]			0·54			0·32
At risk of poverty: < 1,251 €/month [n, %]	178 (44·1%)	181 (43·0%)		139 (50·7%)	120 (44·8%)	
Not at risk of poverty: >=1,251 - 4,169 €/month [n, %]	143 (35·4%)	147 (34·9%)		95 (34·7%)	99 (36·9%)	
Declined to answer [n, %]	73 (18·1%)	81 (19·2%)		35 (12·8%)	44 (16·4%)	
Missing [n, %]	10 (2·5%)	12 (2·9%)		5 (1·8%)	5 (1·9%)	

[a] Chi-Square test for categorical variables or chi-square test with Yates' continuity correction for 2 × 2 tables. Without the categories "Declined to answer" and "Missing".

[d] the mean net equivalent income in Germany in 2021 was 2,424 Euro/month, the respective median was 2,085 Euro/month (Statistisches Bundesamt 2022). The cut-off for the risk of poverty is defined as 60% of the median (€ 2,084·58 * 0·6 = € 1,250·75; Bundesministerium für Arbeit und Soziales 2022, p. 525).

Table S4) Participants' characteristics at baseline in children/adolescents and adults

	Children/adolescents			Adults		
	Standard care cohort (N=42)	**Innovative care cohort (N=25)**	***p* value[a]**	**Standard care cohort (N=636)**	**Innovative care cohort (N=664)**	***p* value[a]**
Age (years) [median, IQR]	16 (14-17)	15 (15-16)	0·76	49 (35-58)	48 (34-57)	0·49
Female sex [n, %]	19 (45·2%)	14 (56·0%)	0·55	385 (60·5%)	407 (61·3%)	0·79
Migratory background			1·00			0·62
No migratory background [n, %]	23 (54·8%)	14 (56·0%)		451 (70·9%)	481 (72·4%)	
Migratory background [n, %]	11 (26·2%)	7 (28·0%)		174 (27·4%)	173 (26·1%)	
Declined to answer [n, %]	1 (2·4%)	0 (0·0%)		11 (1·7%)	10 (1·5%)	
Missing [n, %]	7 (16·7%)	4 (16·0%)		0 (0·0%)	0 (0·0%)	
Highest school education			0·73			0·71
No graduation [n, %]	39 (92·9%)	24 (96·0%)		16 (2·5%)	14 (2·1%)	
Lower secondary (ISCED 2)[b] [n, %]	1 (2·4%)	0 (0·0%)		361 (56·8%)	364 (54·8%)	
Upper secondary (ISCED 3)[b] [n, %]	2 (4·8%)	1 (4·0%)		244 (38·4%)	266 (40·1%)	
Other educational degree [n, %]	0 (0·0%)	0 (0·0%)		3 (0·5%)	5 (0·8%)	
Missing [n, %]	0 (0·0%)	0 (0·0%)		12 (1·9%)	15 (2·3%)	
Highest post-secondary education			1·00			0·13
Currently enrolled in secondary/ tertiary/vocational education [n, %]	40 (95·2%)	24 (96·0%)		52 (8·2%)	43 (6·5%)	
No tertiary/vocational education, not currently enrolled [n, %]	1 (2·4%)	1 (4·0%)		47 (7·4%)	49 (7·4%)	
Vocational qualification (ISCED 4)[b] [n, %]	0 (0·0%)	0 (0·0%)		329 (51·7%)	315 (47·4%)	
Bachelor's/postgraduate degree (ISCED 5-8)[b] [n, %]	0 (0·0%)	0 (0·0%)		188 (29·6%)	234 (35·2%)	
Other tertiary/vocational degree [n, %]	0 (0·0%)	0 (0·0%)		1 (0·2%)	3 (0·5%)	
Declined to answer [n, %]	1 (2·4%)	0 (0·0%)		19 (3·0%)	20 (3·0%)	
Missing [n, %]	0 (0·0%)	0 (0·0%)		0 (0·0%)	0 (0·0%)	
Employment status			0·57			0·077
Full-time [n, %]	0 (0·0%)	0 (0·0%)		205 (32·2%)	231 (34·8%)	
Part-time or less [n, %]	1 (2·4%)	0 (0·0%)		88 (13·8%)	109 (16·4%)	
Unemployed [n, %]	1 (2·4%)	0 (0·0%)		72 (11·3%)	60 (9·0%)	
Retired due to disability [n, %]	0 (0·0%)	0 (0·0%)		91 (14·3%)	67 (10·1%)	
Outside of the labour force for other reasons [n, %]	35 (83·3%)	20 (80·0%)		136 (21·4%)	143 (21·5%)	
Declined to answer [n, %]	5 (11·9%)	5 (20·0%)		44 (6·9%)	54 (8·1%)	
Missing [n, %]	0 (0·0%)	0 (0·0%)		0 (0·0%8	0 (0·0%)	
Duration of main symptom (years) [median, IQR]	8 (2-13)	8 (3-11)	0·84	6 (3-16)	6 (3-16)	0·55
Missing [n, %]	0 (0·0%)	0 (0·0%)		2 (0·3%)	7 (1·1%)	
Number of HPO codes per patient [median, IQR]	5 (3-7)	5 (3-8)	0·66	6 (3-9)	5 (3-8)	0·77
Missing [n, %]	0 (0·0%)	2 (7·0%)		2 (0·7%)	0 (0·0%)	
Disability formally acknowledged [n, %][c]	4 (9·5%)	4 (16·0%)	0·73	245 (38·5%)	251 (37·8%)	0·80
Declined to answer [n, %]	8 (19·0%)	4 (16·0%)		18 (2·8%)	17 (2·6%)	
HrQoL: EQ-5D-VAS [median, IQR]	70 (50-90)	52 (30-67)	0·021	50 (30-70)	50 (30-65)	0·54
Missing [n, %]	0 (0·0%)	0 (0·0%)		0 (0·0%)	1 (0·2%)	

Abbreviations: HPO = human phenotype ontology, HRQoL = health-related quality of life, VAS = visual analogue scale (0-100)

[a] Chi-square test for categorical variables or chi-square test with Yates' continuity correction for 2 × 2 tables. Without the categories "Declined to answer", "Missing" and the open categories "Other tertiary/vocational degree" and "Other educational degree". Non-parametric test (Mann-Whitney *U*-test) for numerical variables.

[b] ISCED – International Standard Classification of Education 2011,

[c] reported by patients

Table S4) Participants' characteristics at baseline in children/adolescents and adults continued

	Children/adolescents			Adults		
	Standard care cohort (N=42)	**Innovative care cohort (N=25)**	***p* value[a]**	**Standard care cohort (N=636)**	**Innovative care cohort (N=664)**	***p* value[a]**
Seen by a mental health specialist in the last 12 months [n,%]	14 (33·3%)	13 (52·0%)	0·16	222 (34·9%)	251 (37·8%)	0·36
Missing [n, %]	0 (0·0%)	1 (4·0%)		9 (1·4%)	4 (0·6%)	
Born outside Germany [n, %]	2 (4·8%)	1 (4·0%)	1·00	87 (13·7%)	72 (10·8%)	0·13
Declined to answer [n, %]	0 (0·0%)	0 (0·0%)		2 (0·3%)	0 (0·0%)	
Missing [n, %]	7 (16·7%)	4 (16·0%)		0 (0·0%)	0 (0·0%)	
Area of residence [n, %]			1·00			0·11
Rural [n, %]	17 (40·5%)	10 (40·0%)		317 (49·8%)	361 (54·4%)	
Urban [n, %]	17 (40·5%)	11 (44·0%)		315 (49·5%)	298 (44·9%)	
Declined to answer [n, %]	1 (2·4%)	0 (0·0%)		0 (0·0%)	0 (0·0%)	
Missing [n, %]	7 (16·7%)	4 (16·0%)		4 (0·6%)	5 (0·8%)	
Assistance in everyday life [n, %]			1·00			0·93
Do not receive assistance [n, %]	32 (76·2%)	19 (76·0%)		543 (85·4%)	572 (86·1%)	
Receive assistance [n, %]	2 (4·8%)	2 (8·0%)		83 (13·1%)	85 (12·8%)	
Declined to answer [n, %]	1 (2·4%)	0 (0·0%)		10 (1·6%)	7 (1·1%)	
Missing [n, %]	7 (16·7%)	4 (16·0%)		0 (0·0%)	0 (0·0%)	
Equivalised disposable income[d] [n, %]			1·00			0·59
At risk of poverty: < 1,251 €/month [n, %]	20 (47·6%)	11 (44·0%)		297 (46·7%)	290 (43·7%)	
Not at risk of poverty: >=1,251 - 4,169 €/month [n, %]	10 (23·8%)	6 (24·0%)		228 (35·8%)	240 (36·1%)	
Declined to answer [n, %]	5 (11·9%)	4 (16·0%)		103 (16·2%)	121 (18·2%)	
Missing [n, %]	7 (16·7%)	4 (16·0%)		8 (1·3%)	13 (2·0%)	

[a] Chi-Square test for categorical variables or chi-square test with Yates' continuity correction for 2 × 2 tables. Without the categories "Declined to answer" and "Missing".

[d] the mean net equivalent income in Germany in 2021 was 2,424 Euro/month, the respective median was 2,085 Euro/month (Statistisches Bundesamt 2022). The cut-off for the risk of poverty is defined as 60% of the median (€ 2,084·58 * 0·6 = € 1,250·75; Bundesministerium für Arbeit und Soziales 2022, p. 525).

Table S5) Most frequently coded symptoms based on Human Phenotype Ontology (HPO)

No.	HPO Code	Symptom	Standard care cohort (N=673)	Innovative care cohort (N=687)	*p* value[a]
1	HP:0003326	Myalgia	111	146	0·030
2	HP:0012378	Fatigue	126	128	0·98
3	HP:0002829	Arthralgia	113	135	0·20
4	HP:0002315	Headache	102	112	0·61
5	HP:0002027	Abdominal pain	94	118	0·12
6	HP:0002321	Dizziness/ vertigo	109	90	0·12
7	HP:0003418	Back pain	110	80	0·015
8	HP:0012531	Non specified pain	109	77	0·0094
9	HP:0003401	Paraesthesia	82	101	0·20
10	HP:0002014	Diarrhoea	66	72	0·75

[a] Chi-square test with Yates' continuity correction for 2 × 2 tables.

Additional study results

Table S6) Diagnoses newly established during the ZSE-DUO project coded by the *International Statistical Classification of Diseases and Related Health Problems* ICD-10 German modification in the standard care and the innovative care cohort.

Note: the table shows the total number of diagnoses established, not the number of patients with a respective diagnosis· In other words, there may be more than one Diagnosis from an ICD-10GM diagnostic category per patient. For a better overview, the codes are primarily grouped by the first letter and then by the first number.

		Total no. of diagnoses		**ICD-10 GM-Groups**																			
		SC	IC	SC	IC	SC	IC	SC	IC	SC	IC	SC	IC	SC	IC	SC	IC	SC	IC	SC	IC	SC	IC
A	Certain infectious and parasitic diseases	0	1	A6x.-																			
				0	1																		
B	Certain infectious and parasitic diseases	2	3	B1x.-		B4x.-		B6x.-		B7x.-		B8x.-											
				1	0	1	0	0	1	0	1	0	1										
C	Neoplasms (C0x·-C9x·-)	1	8	C3x.-		C4x.-		C5x.-		C6x.-		C8x·-		C9x.-									
				0	1	0	1	0	2	0	2	1	1	0	1								
D	Neoplasms (D0x·-D4x·-) Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (D5x·-D9x·-)	19	28	D1x.-		D2x.-		D3x.-		D4x.-		D5x.-		D6x.-		D7x.-		D8x.-					
				3	2	0	1	0	1	6	5	1	4	2	2	2	5	5	8				
E	Endocrine, nutritional and metabolic diseases	24	39	E0x.-		E1x.-		E2x.-		E5x.-		E6x.-		E7x.-		E8x.-							
				0	3	0	1	4	1	10	13	0	5	5	8	5	8						
F	Mental and behavioural disorders	45	686	F0x.-		F1x.-		F2x.-		F3x.-		F4x.-		F5x.-		F6x.-		F7x.-		F8x.-		F9x.-	
				2	10	1	30	0	4	10	144	26	404	2	63	1	16	0	5	1	4	2	6
G	Diseases of the nervous system	105	186	G0x.-		G1x.-		G2x.-		G3x.-		G4x.-		G5x.-		G6x.-		G7x.-		G8x.-		G9x.-	
				2	2	9	27	10	25	6	6	12	20	4	13	41	56	9	7	0	7	12	23
H	Diseases of the eye and adnexa (H0x·-H5x·-) Diseases of the ear and mastoid process (H6x·-H9x·-)	12	3	H2x.-		H3x.-		H5x.-		H8x.-		H9x.-											
				2	0	7	0	3	0	0	1	0	2										
I	Diseases of the circulatory system	4	23	I1x.-		I2x.-		I3x.-		I4x.-		I5x.-		I6x.-		I7x.-		I8x.-		I9x.-			
				1	2	0	1	0	2	1	3	0	1	0	4	1	6	1	1	0	3		
J	Diseases of the respiratory system	2	4	J3x.-																			
				2	4																		
K	Diseases of the digestive system	15	13	K0x.-		K1x.-		K2x.-		K5x.-		K6x.-		K7x.-		K9x.-							
				2	0	4	3	2	0	3	7	2	1	1	1	1	1						
L	Diseases of the skin and subcutaneous tissue	12	24	L0x.-		L1x.-		L2x.-		L3x.-		L4x.-		L5x.-		L6x.-		L7x.-		L8x.-		L9x.-	
				0	1	1	0	2	1	0	2	1	4	3	4	1	2	2	4	0	4	2	2
M	Diseases of the musculoskeletal system and connective tissue	25	78	M0x.-		M1x.-		M2x.-		M3x.-		M4x.-		M5x.-		M6x.-		M7x.-		M8x.-			
				0	4	0	2	0	6	14	45	3	4	3	0	0	2	3	12	2	3		
N	Diseases of the genitourinary system	1	6	N1x.-		N3x.-		N4x.-		N8x.-		N9x.-											
				0	2	0	1	0	1	1	1	0	1										
Q	Congenital malformations, deformations and chromosomal abnormalities	13	20	Q0x.-		Q3x.-		Q4x.-		Q6x.-		Q7x.-		Q8x.-		Q9x.-							
				0	1	2	0	0	1	2	2	5	6	3	8	1	2						
R	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	3	28	R0x.-		R1x.-		R2x.-		R4x.-		R5x.-		R6x.-		R7x.-		R9x.-					
				0	4	0	3	2	8	0	4	0	4	0	3	1	1	0	1				

T	Injury, poisoning and certain other consequences of external causes	5	2	T6x.-		T7x.-		T8x.-						
				1	0	4	1	0	1					
Y	External causes of morbidity and mortality	0	1	Y5x.-										
				0	1									
Z	Factors influencing health status and contact with health services	3	5	Z0x.-		Z5x.-		Z7x.-		Z8x.-				
				1	0	0	1	0	3	2	1			

Patients with diagnoses explaining the entire symptomatic spectrum presented

Table S7) Proportion of patients with diagnoses fully explaining the symptomatic spectrum irrespective of the type of diagnoses (primary outcome).
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with newly confirmed diagnoses explaining the entire symptomatic spectrum presented. In addition to the basic model, more complex models including covariates and interactions are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from model 3 due to multicollinearity with SC/IC. Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			Covariates Models					
	Model 1			Model 2			Model 3		
	OR	95% CrI		OR	95% CrI		OR	CrI[a]	
Fixed part									
SC/IC	3·45*	1·99	5·65	3·62*	2·05	6·02	4·11*	2·16	7·22
Sex				1·07	0·81	1·38	1·02	0·70	1·42
Age				1·01*	1·00	1·02	1·01*	1·00	1·02
Diploma[b]				0·95	0·70	1·24	1·28	0·78	1·97
Student[c]				1·24	0·75	1·94	1·42	0·71	2·53
Interactions									
SC/IC x Sex							1·13	0·76	1·61
SC/IC x Diploma[b]							0·66	0·39	1·04
SC/IC x Student[c]							0·84	0·40	1·52
Random part: Variances									
Level 3 – CRD	0·26*	0·01	0·92	0·33	0·01	1·09	0·32	0·01	1·06
Level 2 – SC/IC	0·29*	0·04	0·80	0·28*	0·03	0·85	0·29*	0·04	0·87
Model fit									
DIC	1504·3			1459·2			1460·1		
BF[d]	319·3			56·9			0·0		
N	1,358			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95% credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In model 3, 90% credibility intervals (90% CrI)) are shown for interactions, while 95% credibility intervals are provided for all other effects.
[b] – Bachelor's/postgraduate degree (ISCED 5-8).
[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model.
The most complex model 3, including interactions, did not result in a better model fit compared to model 2.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S8) Proportion of patients with diagnoses fully explaining the symptomatic spectrum, including at least one rare disease diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with newly confirmed diagnoses explaining the entire symptomatic spectrum presented. In addition to the basic model, more complex models including covariates and interactions are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from models 3 and 3b due to multicollinearity with SC/IC Despite a high thinning rate and a high simulation rate, model 3 showed significant auto-correlation. Therefore, non-significant interaction terms were eliminated from model 3 (model 3b). Analyses were performed using Stata 15.1. Data are presented as odds-ratios, their standard deviations, and credibility intervals.

	Basic Model			**Covariates Models**								
	Model 1			**Model 2**			**Model 3**			**Model 3b**		
	OR	**95% CrI**		**OR**	**95% CrI**		**OR**	**CrI[a]**		**OR**	**CrI[a]**	
Fixed part												
SC/IC	1·31	0·73	2·13	1·39	0·77	2·30	2·50*	1·21	4·54	1·78	0·93	3·03
Sex				1·25	0·83	1·83	1·72	0·95	2·88	1·25	0·83	1·82
Age				1·02*	1·00	1·03	1·01*	1·00	1·03	1·02*	1·00	1·03
Diploma[b]				0·97	0·63	1·45	1·61	0·83	2·78	1·57	0·81	2·70
Student[c]				1·11	0·47	2·14	1·17	0·37	2·56	1·11	0·48	2·16
Interactions												
SC/IC x Sex							0·64	0·33	1·07			
SC/IC x Diploma[b]							0·45*	0·21	0·83	0·49*	0·23	0·88
SC/IC x Student[c]							1·01	0·26	2·42			
Random part: Variances												
Level 3 – CRD	0·33	0·01	1·17	0·36	0·01	1·26	0·37	0·02	1·26	0·37	0·01	1·25
Level 2 – SC/IC	0·19*	0·01	0·79	0·20*	0·01	0·87	0·19	0·01	0·77	0·19*	0·01	0·77
Model fit												
DIC	859·8			833·2			831·7			831,0		
BF[d]	0·1			12·1			0·0			0·002		
N	1,358			1,319			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95%-credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In models 3 and 3b, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects.
[b] – Bachelor's/postgraduate degree (ISCED 5-8).
[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model. Model 3b which includes only significant interaction terms showed the best modelfit.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S9) Proportion of patients with diagnoses fully explaining the symptomatic spectrum, including at least one mental disorder diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with newly confirmed diagnoses explaining the entire symptomatic spectrum presented. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from models 3 and 3b due to multicollinearity with SC/IC. Despite a high thinning rate and a high simulation rate, model 3 showed significant auto-correlation. Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			**Covariates Models**								
	Model 1			**Model 2**			**Model 3**			**Model 3b**		
	OR	**95% CrI**		**OR**	**95% CrI**		**OR**	**CrI[a]**		**OR**	**CrI[a]**	
Fixed part												
SC/IC	16·98*	9·14	29,78	17·06*	9·12	29·74	10·46*	4·15	23·23	10·82*	4·79	22·25
Sex				1·11	0·77	1·56	0·56	0·24	1·14	0·53	0·17	1·22
Age				1·00	0·99	1·02	1·00	0·99	1·02	1·00	0·99	1·02
Diploma[b]				0·77	0·52	1·10	0·60	0·26	1·17	0·77	0·52	1·10
Student[c]				1·05	0·52	1·84	1·07	0·19	3·07	1·06	0·53	1·89
Interactions												
SC/IC x Sex							2·61*	1·22	4·70	3·04*	1·13	6·34
SC/IC x Diploma[b]							1·52	0·75	2·69			
SC/IC x Student[c]							1·65	0·36	4·43			
Random part: Variances												
Level 3 – CRD	0·76	0·09	2·25	0·81	0·08	2·38	0·78	0·06	2·35	0·81	0·10	2·36
Level 2 – SC/IC	0·11*	0·01	0·55	0·12*	0·01	0·66	0·13*	0·01	0·72	0·12*	0·01	0·59
Model fit												
DIC	936·0			913·6			913·6			912·0		
BF[d]	>150			27·1			0·0			0·0		
N	1,358			1,319			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95% credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In models 3 and 3b, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects.

[b] – Bachelor's/postgraduate degree (ISCED 5-8).
[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model. Model 3b which includes only significant interaction terms showed the best model fit.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S10) Proportion of patients with diagnoses fully explaining the symptomatic spectrum, including at least one non-rare somatic disease diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with newly confirmed diagnoses explaining the entire symptomatic spectrum presented. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from models 3 and 3b due to multicollinearity with SC/IC. Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			**Covariates Models**					
	Model 1			**Model 2**			**Model 3**		
	OR	**95% CrI**		**OR**	**95% CrI**		**OR**	**CrI[a]**	
Fixed part									
SC/IC	2·26*	1·50	3·27	2·30*	1·51	3·36	2·52*	1·31	4·50
Sex				1·01	0·71	1·38	0·89	0·52	1·43
Age				1·01*	1·00	1·02	1·01*	1·00	1·02
Diploma[b]				1·08	0·74	1·51	1·48	0·97	2·18
Student[c]				1·45	0·76	2·53	1·83	0·76	3·72
Interactions									
SC/IC x Sex							1·32	0·76	2·07
SC/IC x Diploma[b]							0·66	0·36	1·08
SC/IC x Student[c]							0·81	0·32	1·60
Random part: Variances									
Level 3 – CRD	0·69	0·16	1·87	0·70	0·16	1·94	0·70	0·16	1·97
Level 2 – SC/IC	0·06*	0·01	0·26	0·06*	0·01	0·34	0·07*	0·01	0·32
Model fit									
DIC	1,036·1			1,021·6			1,460·1		
BF[d]	28·8			0·0			0·0		
N	1,358			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI– 95% credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In model 3, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects.
[b] – Bachelor's/postgraduate degree (ISCED 5-8).

[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model. Model 2, excluding non-significant interaction terms, showed the best model fit.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S11) Proportion of female participants with diagnoses fully explaining the symptomatic spectrum.

	Standard care cohort (N=401)	Innovative care cohort (N=418)	*p* value[a]
All female participants with explaining diagnoses [n,%]	74 (18·5%)	176 (42·1%)	<0·0001
Female participants with explaining diagnoses and contribution of at least one rare disease [n,%]	40 (10·0%)	47 (11·2%)	0·63
Female participants with explaining diagnoses and contribution of at least one mental disorder [n,%]	8 (2·0%)	125 (29·9%)	<0·0001
Female participants with explaining diagnoses and contribution of at least one non-rare disease [n,%]	37 (9·2%)	79 (18·9%)	0·00011

[a] Chi-square test with Yates' continuity correction for 2 × 2 tables.

Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

Table S12) Proportion of male participants with diagnoses fully explaining the symptomatic spectrum.

	Standard care cohort (N=401)	Innovative care cohort (N=418)	*p* value[a]
All male participants with explaining diagnoses [n,%]	52 (19·2%)	109 (40·7%)	<0·0001
Male participants with explaining diagnoses and contribution of at least one rare disease [n,%]	19 (7·0%)	30 (11·2%)	0·12
Male participants with explaining diagnoses and contribution of at least one mental disorder [n,%]	11 (4·1%)	71 (26·5%)	<0·0001
Male participants with explaining diagnoses and contribution of at least one non-rare disease [n,%]	29 (10·7%)	50 (18·7%)	0·013

[a] Chi-square test with Yates' continuity correction for 2 × 2 tables.

Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

Table S13) Proportion of children/adolescents with diagnoses fully explaining the symptomatic spectrum.

	Standard care cohort (N=41)	Innovative care cohort (N=24)	*p* value[a]
All children/Adolescents with explaining diagnoses [n,%]	14 (34·2%)	9 (37·5%)	1·0
Children/Adolescents with explaining diagnoses and contribution of at least one rare disease [n,%]	6 (14·6%)	2 (8·3%)	0·72
Children/Adolescents with explaining diagnoses and contribution of at least one mental disorder [n,%]	1 (2·4%)	6 (25·0%)	0·016
Children/Adolescents with explaining diagnoses and contribution of at least one non-rare disease [n,%]	9 (22·0%)	3 (12·5%)	0·54

[a] Chi-square test with Yates' continuity correction for 2 × 2 tables.

Please note: statistical analyses are limited by few cases per cell. Some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

Table S14) Proportion of adults with diagnoses fully explaining the symptomatic spectrum.

	Standard care cohort (N=631)	Innovative care cohort (N=662)	*p* value[a]
All adults with explaining diagnoses [n,%]	112 (17·7%)	276 (41·7%)	<0·0001
Adults with explaining diagnoses and contribution of at least one rare disease [n,%]	53 (8·4%)	75 (11·3%)	0·095
Adults with explaining diagnoses and contribution of at least one mental disorder [n,%]	18 (2·9%)	190 (28·7%)	<0·0001
Adults with explaining diagnoses and contribution of at least one non-rare disease [n,%]	57 (9·0%)	126 (19·0%)	<0·0001

[a] Chi-square test with Yates' continuity correction for 2 × 2 tables.

Please note: some patients had diagnoses from more than one diagnostic category and are included in each applicable category.

Table S15) Time to diagnoses fully explaining the symptomatic spectrum irrespective of the type of diagnoses (Mann-Whitney *U*-test).
Mann-Whitney *U*-test comparing standard care and innovative care with respect to the time to newly confirmed diagnoses explaining the entire symptomatic spectrum presented. Analyses were performed using Stata and SPSS 28.0.

	SC			IC			Mann-Whitney *U*-test			Hodges-Lehman median difference		
	N	***Mdn***	***IQR***	**N**	***Mdn***	***IQR***	***U***	**z**	***p* value**	**Estimate**	**95% CI**	
Total	126	2·5	1·0 - 5·0	285	0·0	0·0 - 4·0	12398	5·203	<0·0001	1·0	1·000	2·000
Rare disease	59	3·0	1·0 - 7·0	77	2·0	1·0 - 6·0	2101	0·755	0·45	0·0	-1·000	1·000
Mental disorder	19	1·0	0·0 - 3·0	196	0·0	0·0 - 2·0	1287	2·518	0·012	1·0	0·000	1·000
Non-rare somatic disease	66	2·0	1·0 - 5·0	129	1·0	0·0 - 4·0	3152	3·053	0·0023	1·0	0·000	2·000
Non-rare disease (Diagnosed by exclusion)	67	2·0	1·0 - 4·0	208	0·0	0·0 - 2·0	4325	5·059	<0·0001	1·0	1·000	2·000

Abbreviations: Mdn – median, IQR – Interquartile range, 95% CI – 95% confidence interval, SC – standard care, IC – innovative care.

Table S16) Time to diagnoses fully explaining the symptomatic spectrum irrespective of the type of diagnoses (linear regression model).
Linear regression models statistically comparing standard care and innovative care with respect to the time to newly confirmed diagnoses explaining the entire symptomatic spectrum presented. Analyses are based on linear regression with bootstrap (repl· 5,000), bias-corrected and accelerated robust standard errors and confidence intervals to counterbalance distortion of the distribution. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses were performed using Stata 15.1.

	Basic Model					**Covariates Models**								
	Model 1					**Model 2**					**Model 3**			
	ß	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**	
SC/IC	-1·37*	0·35	0·00020	-2·08	-0·70	-1·41*	0·36	<0·0001	-2·15	-0·73	-0·51	0·69	0·46	-1·90	0·86
Sex						0·27	0·32	0·41	-0·38	0·88	0·98	0·63	0·12	-0·32	2·19
Age						0·00	0·01	0·93	-0·02	0·02	0·00	0·01	0·92	-0·02	0·02
Diploma[a]						0·26	0·37	0·48	-0·42	1·03	0·89	0·72	0·21	-0·48	2·30
Student[b]						0·01	0·56	0·99	-1·02	1·17	0·08	0·93	0·93	-1·59	2·08
Interactions															
SC/IC x Sex											-1·04	0·75	0·17	-2·51	0·44
SC/IC x Diploma[a]											-0·90	0·84	0·28	-2·50	0·75
SC/IC x Student[b]											-0·03	1·08	0·98	-2·22	1·98
Model fit															
Wald χ^2(df), *p*	14·91*	1	0·00073			15·55*	5	0·0082			17·87*	8	0·022		
R^2	0·04					0·04					0·05				
R^2 adj·	0·04					0·03					0·03				
Model comparison															
Likelood Ratio χ^2						52*		<0·0001			3·21		0·36		
AIC	2,121					2,077					2,080				
BIC	2,129					2,101					2,116				
N	411					400					400				

Abbreviations: ß – regression coefficient, SE – standard error, 95% CI – 95% confidence interval, SC – standard care, IC – innovative care.
[a] – Bachelor's/postgraduate degree (ISCED 5-8).
[b] – Currently enrolled in secondary/tertiary/ vocational education.
* - p <0·05.
Model 2, excluding interaction terms, showed the best model fit.

Table S17) Successful referral of patients with explaining diagnoses to local regular care in the standard care and innovative care cohorts relative to the respective entire cohort.
Transferred cases are related to the total size of respective cohort (standard care: n=658; innovative care: n=659).

	Successful Referral SC		IC		Fisher's Exact Test	Effect size
	N	%	N	%	*p* value	Cramer's V
Total	56	8·5%	126	19·1%	<0·0001	0·15
RD	21	3·2%	45	6·8%	0·0034	0·08
MD	13	2·0%	82	12·4%	<0·0001	0·20
Non-RD	32	4·9%	50	7·6%	0·052	0·06

SC – standard care, IC – innovative care, RD – rare disease, MD – mental disorder, non-RD – non-rare somatic disease.

Table S18) Successful referral of patients with explaining diagnoses to local regular care in the standard care and innovative care cohorts relative to the respective sample with at least one newly established diagnosis.
Transferred cases are related to the number of patients with at least one newly established diagnosis in the respective cohort (standard care: n=152; innovative care: n=374).

	Successful Referral SC		IC		Fisher's Exact Test	Effect size
	N	%	N	%	*P* value	Cramer's V
Total	56	36·8%	126	33·7%	0·54	-0·03
RD	21	13·8%	45	12·0%	0·56	-0·02
MD	13	8·6%	82	21·9%	0·00025	0·16
Non RD	32	21·1%	50	13·4%	0·034	0·10

SC – standard care, IC – innovative care, RD – rare disease, MD – mental disorder, non-RD – non-rare somatic disease·

Patients with at least one newly established diagnosis

Table S19) Proportion of patients with any newly established diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with any newly confirmed diagnosis. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from model 3 due to multicollinearity with SC/IC. Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			Covariates Models					
	Model 1			Model 2			Model 3		
	OR	95% CrI		OR	95% CrI		OR	CrI[a]	
Fixed part									
SC/IC	7·91*	3·62	15·48	8·35*	3·70	17·16	9·43*	3·90	20·33
Sex				1·29	0·98	1·66	1·33	0·91	1·90
Age				1·01*	1·00	1·02	1·01*	1·00	1·02
Diploma[b]				0·90	0·67	1·18	1·02	0·66	1·53
Student[c]				0·92	0·56	1·44	0·93	0·53	1·58
Interactions									
SC/IC x Sex							0·99	0·63	1·45
SC/IC x Diploma[b]							0·84	0·51	1·24
SC/IC x Student[c]							1·02	0·44	1·91
Random part: Variances									
Level 3 – CRD	0·58	0·02	2·02	0·64	0·02	2·28	0·63	0·01	2·18
Level 2 – SC/IC	0·56	0·12	1·58	0·62	0·14	1·70	0·63	0·14	1·74
Model fit									
DIC	1,495·7			1,452·5			1,456·3		
BF[d]	0·2			0·0			0·0		
N	1,358			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95%-credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In model 3, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects.

[b] – Bachelor's/postgraduate degree (ISCED 5-8).
[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model.
The most complex model 3, including interactions, did not result in a better model fit compared to model 2.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S20) Proportion of patients with any newly established diagnosis, including at least one rare disease diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with any newly confirmed diagnosis. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from models 3 and 3b due to multicollinearity with SC/IC. Despite a high thinning rate and a high simulation rate, model 3 showed significant auto-correlation. Therefore, non-significant interaction terms were eliminated from the model However, in the new model, there was still a problem with auto-correlation and the SC/IC x sex interaction was not significant any more. Therefore, this interaction term was also removed from the model (model 3b). Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			**Covariates Models**								
	Model 1			**Model 2**			**Model 3**			**Model 3b**		
	OR	**95% CrI**		**OR**	**95% CrI**		**OR**	**CrI[a]**		**OR**	**CrI[a]**	
Fixed part												
SC/IC	1·44	0·89	2·19	1·48	0·88	2·36	3·10*	1·93	4·74	2·04*	1·12	3·47
Sex				1·23	0·84	1·75	1·71*	1·10	2·55	1·24	0·85	1·77
Age				1·01	1·00	1·03	1·01*	1·00	1·03	1·01	1·00	1·03
Diploma[b]				1·04	0·70	1·50	2·00*	1·13	3·36	1·84*	1·00	3·09
Student[c]				0·94	0·42	1·77	1·28	0·43	2·83	0·95	0·43	1·76
Interactions												
SC/IC x Sex							0·59*	0·42	0·80			
SC/IC x Diploma[b]							0·36*	0·18	0·61	0·41*	0·21	0·71
SC/IC x Student[c]							0·73	0·19	1·70			
Random part: Variances												
Level 3 – CRD	0·42	0·03	1·37	0·41	0·02	1·37	0·43	0·02	1·40	0·42	0·02	1·40
Level 2 – SC/IC	0·12*	0·01	0·54	0·16*	0·01	0·69	0·16*	0·01	0·68	0·16*	0·01	0·69
Model fit												
DIC	958·6			929·1			924·0			924·7		
BF[d]	0·1			0·0			0·0			4·7		
N	1,358			1,319			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95%-credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In models 3 and 3b, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects·.
[b] – Bachelor's/postgraduate degree (ISCED 5-8).
[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model.
Model 3b, which includes only significant interaction terms, showed the best model fit.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S21) Proportion of patients with any newly established diagnosis, including at least one mental disorder diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with any newly confirmed diagnosis. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from models 3 and 3b due to multicollinearity with SC/IC. Despite a high thinning rate and a high simulation rate, model 3 showed significant auto-correlation. Therefore, non-significant interaction terms were eliminated from the model (model 3b). Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			**Covariates Models**								
	Model 1			**Model 2**			**Model 3**			**Model 3b**		
	OR	**95% CrI**		**OR**	**95% CrI**		**OR**	**CrI[a]**		**OR**	**CrI[a]**	
Fixed part												
SC/IC	33·32*	15·98	65·38	32·72*	15·13	66·64	18·12*	8·29	35·55	21·82*	8·90	49·16
Sex				1·17	0·85	1·58	0·56*	0·31	0·94	0·63	0·26	1·29
Age				1·00	0·99	1·01	1·00	0·99	1·01	1·00	0·99	1·01
Diploma[b]				0·87	0·61	1·19	0·71	0·23	1·56	0·88	0·62	1·21
Student[c]				0·94	0·50	1·61	0·80	0·24	1·87	0·94	0·50	1·59
Interactions												
SC/IC x Sex							2·55*	1·53	3·91	2·47*	1·08	4·64
SC/IC x Diploma[b]							1·62	0·61	3·42			
SC/IC x Student[c]							1·58	0·53	3·52			
Random part: Variances												
Level 3 – CRD	0·48	0·02	1·58	0·47	0·01	1·65	0·46	0·01	1·55	0·48	0·01	1·63
Level 2 – SC/IC	0·36	0·02	1·16	0·42	0·02	1·51	0·38	0·02	1·25	0·39	0·02	1·35
Model fit												
DIC	1,106·0			1,086·5			1,086·1			1,084·7		
BF[d]	>150			0·0			0·0			0·0		
N	1,358			1,319			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95%-credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In models 3 and 3b, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects.
[b] – Bachelor's/postgraduate degree (ISCED 5-8).
[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model, for model 3b to model 2.
Model 3b, which includes only significant interaction terms, showed the best model fit.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S22) Proportion of patients with any newly established diagnosis, including at least one non-rare somatic disease diagnosis.
Mixed logistic regression models statistically comparing standard care and innovative care with respect to the proportion of patients with any newly confirmed diagnosis. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses are based on Bayesian inference with uninformative priors and simulations (500,000-1,000,000) employing Markov Chain Monte Carlo sampling. The interaction effect for SC/IC*age was excluded from model 3 due to multicollinearity with SC/IC. Analyses were performed using Stata 15.1. Data are presented as odds-ratios, standard deviations and credibility intervals.

	Basic Model			**Covariates Models**					
	Model 1			**Model 2**			**Model 3**		
	OR	**95% CrI**		**OR**	**95% CrI**		**OR**	**CrI[a]**	
Fixed part									
SC/IC	2·25*	1·28	3·79	2·29*	1·25	3·95	1·77	0·95	2·94
Sex				1·25	0·94	1·65	1·19	0·78	1·74
Age				1·01*	1·00	1·02	1·01*	1·00	1·02
Diploma[b]				0·90	0·65	1·20	0·71	0·49	0·97
Student[c]				0·99	0·57	1·60	0·94	0·51	1·55
Interactions									
SC/IC x Sex							1·16	0·72	1·74
SC/IC x Diploma[b]							1·39	0·95	2·00
SC/IC x Student[c]							1·34	0·59	2·53
Random part: Variances									
Level 3 – CRD	0·94	0·12	2·71	0·99	0·09	2·88	0·97	0·07	2·81
Level 2 – SC/IC	0·27*	0·03	0·91	0·31	0·05	1·07	0·31	0·04	1·07
Model fit									
DIC	1,348·5			1,319·4			1,323·2		
BF[d]	>150			0·0			0·0		
N	1,358			1,319			1,319		

Abbreviations: OR – odds ratio, SD – standard deviation, 95% CrI – 95%-credibility interval, CrI – credibility interval, SC – standard care, IC – innovative care, CRD – centre for rare diseases, DIC – Deviance Information Criterion, BF – Bayes Factors.
[a] – In model 3, 90% credibility intervals (90% CrI) are shown for interactions, while 95% credibility intervals are provided for all other effects.
[b] – Bachelor's/postgraduate degree (ISCED 5-8).

[c] – Currently enrolled in secondary/tertiary/ vocational education.
[d] – Model comparison to the previous model, for the basic model comparison to an empty model.
Model 2, excluding non-significant interaction terms, showed the best model fit.
Main effects *p <0·05 (95% CrI), interactions *p<0·1 (90% CrI).

Table S23) Time to first newly established diagnosis in patients with at least one newly established diagnosis (Mann-Whitney *U*-test).
Mann-Whitney *U*-test comparing standard care and innovative care with respect to the time to newly confirmed diagnoses explaining the entire symptomatic spectrum presented. Analyses were performed using Stata SPSS 28.0.

	SC			IC			Mann-Whitney *U*-test			Hodges-Lehman median difference		
	N	***Mdn***	***IQR***	**N**	***Mdn***	***IQR***	***U***	**z**	***p* value**	**Estimate**	**95% CI**	
Total	201	3·0	1·0 - 5·0	486	0·0	0·0 - 4·0	31956	7·520	<0·0001	1·0	1·000	2·000
Rare disease	67	3·0	1·0 - 7·0	97	3·0	1·0 - 6·0	2979	0·588	0·56	0·0	-1·000	1·000
Mental disorder	28	1·0	0·0 - 3·8	346	0·0	0·0 - 2·0	3468	2·973	0·0029	1·0	0·000	1·000
Non-rare somatic disease	127	3·0	1·0 - 5·0	209	1·0	0·0 - 5·0	10036	3·848	0·00012	1·0	0·000	2·000

Abbreviations: Mdn – median, IQR – Interquartile range, 95% CI – 95% confidence interval, SC – standard care, IC – innovative care.

Table S24) Time to first newly established diagnosis in patients with at least one newly established diagnosis (linear regression model).
Linear regression models statistically comparing standard care and innovative care with respect to the time to first newly confirmed diagnosis. Analyses are based on linear regression with bootstrap (repl· 5,000), bias-corrected and accelerated robust standard errors and confidence intervals to counterbalance distortion of the distribution. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses were performed using Stata 15.1. Data are presented as correlation coefficients, standard errors and confidence intervals.

	Basic Model					**Covariates Models**									
	Model 1					**Model 2**					**Model 3**				
	ß	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**	
SC/IC	-1·49*	0·28	0·00020	-2·05	-0·96	-1·50*	0·28	0·00020	-2·07	-0·98	-0·61	0·54	0·26	-1·69	0·41
Sex						0·40	0·26	0·13	-0·12	0·90	1·13*	0·49	0·022	0·15	2·10
Age						0·01	0·01	0·55	-0·01	0·02	0·01	0·01	0·58	-0·01	0·02
Diploma[a]						0·20	0·29	0·50	-0·34	0·82	0·70	0·59	0·24	-0·41	1·87
Student[b]						0·06	0·45	0·89	-0·78	1·00	0·37	0·68	0·58	-0·88	1·77
Interactions															
SC/IC x Sex											-1·02	0·59	0·082	-2·18	0·17
SC/IC x Diploma[a]											-0·68	0·68	0·32	-2·03	0·62
SC/IC x Student[b]											-0·44	0·79	0·57	-2·01	1·12
Model fit															
Wald χ^2 (df) *p*	29·07*	1	<0·0001			30·91*	5	<0·0001			33·47*	8	<0·0001		
R^2	0·04					0·05					0·05				
R^2 adj·	0·04					0·04					0·04				
Model comparison															
Likelood Ratio χ^2						97·73*		<0·0001			3·78		0·286		
AIC	3,579					3,489					3,491				
BIC	3,588					3,516					3,352				
N	687					667					667				

Abbreviations: ß – regression coefficient, SE – standard error, 95% CI – 95% confidence interval, SC – standard care, IC – innovative care.
[a] - Bachelor's/postgraduate degree (ISCED 5-8).
[b] - Currently enrolled in secondary/tertiary/ vocational education.
* - p <0·05.
Model 2, excluding non-significant interaction terms, showed the best model fit.

Table S25) Successful referral of patients with any new diagnosis to local regular care in the standard care and innovative care cohorts relative to the respective entire cohort (Fisher's exact test).
Transferred cases are related to the total size of the respective cohort (standard care: N=658; innovative care: N=659).

	Successful Referral				Fisher's Exact Test	Effect size
	SC		IC			
	N	%	N	%	*p* value	Cramer's V
Total	81	12·3%	181	27·5%	<0·0001	0·19
RD	26	4·0%	48	7·3%	0·012	0·07
MD	15	2·3%	124	18·8%	<0·0001	0·27
Non-RD	51	7·8%	73	11·1%	0·047	0·06

SC – standard care, IC – innovative care, RD – rare disease, MD – mental disorder, non-RD – non-rare somatic disease.

Table S26) Successful referral of patients with any new diagnosis to local regular care in the standard care and innovative care cohorts relative to the respective entire cohort (logistic regression model).

Logistic regression models statistically comparing standard care and innovative care with respect to the successful referral of patients with any newly confirmed diagnosis. In addition to the basic model, more complex models, including covariates and interactions, are presented. The interaction effect for SC/IC*age was excluded from model 3 due to multicollinearity with SC/IC. Analyses were performed using SPSS 28.0. Data are presented as odds-ratios, standard errors and confidence intervals.

	Basic Model					**Covariates Models**									
	Model 1					**Model 2**					**Model 3**				
	OR	**SE**	***p* value**	**95% CI**		**OR**	**SE**	***p* value**	**95% CI**		**OR**	**SE**	***p* value**	**95% CI**	
SC/IC	2·70*	0·15	<0·0001	2·02	3·60	2·66*	0·15	<0·0001	1·98	3·56	3·17*	0·27	<0·0001	1·86	5·41
Sex						0·85	0·15	0·28	0·64	1·13	0·92	0·24	0·74	0·57	1·49
Age						1·00	0·01	0·81	0·99	1·01	1·00	0·01	0·81	0·99	1·01
Diploma[a]						0·82	0·16	0·23	0·60	1·13	1·12	0·27	0·67	0·67	1·89
Student[b]						0·84	0·27	0·53	0·49	1·44	0·69	0·42	0·39	0·30	1·59
Interactions															
SC/IC x Sex											0·89	0·30	0·69	0·49	1·61
SC/IC x Diploma[a]											0·63	0·33	0·16	0·33	1·20
SC/IC x Student[b]											1·39	0·50	0·51	0·52	3·74
Model fit		df					df					df			
Wald χ^2	48·46*	1	<0·0001			47·93*	5	<0·0001			51·01*	8	<0·0001		
Nagelskerger Pseudo-R^2	0·06					0·06					0·06				
Hosmer-Lemeshow χ^2	0·00	0·00	-			8·43	8	0·39			4·63	8	0·80		
AUC	0·62*	0·02	<0·0001	0·58	0·66	0·63*	0·02	<0·0001	0·60	0·67	0·64*	0·02	<0·0001	0·60	0·68
N	1317					1278					1278				

Abbreviations: OR – odds ratio, SE – standard error, 95% CI – 95% confidence interval, SC – standard care, IC – innovative care.

[a] - Bachelor's/postgraduate degree (ISCED 5-8).

[b] - Currently enrolled in secondary/tertiary/ vocational education·

* - p <0·05.

Model 3 showed the best model fit.

Table S27) Successful referral of patients with any new diagnosis to local regular care in the standard care and innovative care cohorts relative to the respective sample with at least one newly established diagnosis.
Transferred cases are related to the number of patients with at least one newly established diagnosis in the respective cohort (standard care: n=152; innovative care: n=374).

	Successful Referral				Fisher's Exact Test	Effect size
	SC		IC			
	N	%	N	%	*p* value	Cramer's V
Total	81	53·3%	181	48·4%	0·34	-0·04
RD	26	17·1%	48	12·8%	0·21	-0·06
MD	15	9·9%	124	33·2%	<0·0001	0·24
Non-RD	51	33·6%	73	19·5%	0·0010	0·15

SC – standard care, IC – innovative care, RD – rare disease, MD – mental disorder, non-RD – non-rare somatic disease.

Patient-reported outcome measures

Table S28) Change in EQ-5D visual analogue scale rating between baseline and 12-month follow-up.
Linear regression models statistically comparing standard care and innovative care. Analyses are based on linear regression with bootstrap (repl· 5,000), bias-corrected and accelerated robust standard errors and confidence intervals to counterbalance distortion of the distribution. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses were performed using SPSS 27.0. Data are presented as correlation coefficients, standard errors and confidence intervals.

	Basic Model					**Covariates Models**									
	Model 1					**Model 2**					**Model 3[$]**				
	ß	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**	
SC/IC	2·47	1·45	0·085	-0·33	5·35	2·51	1·45	0·081	-0·28	5·33	-·64	2·85[$]	0·82[$]	-6·30[$]	4·02[$]
Sex						1·29	1·50	0·38	-1·66	4·20	-1·60	1·96[$]	0·42[$]	-5·37[$]	2·15[$]
Age						0·02	0·05	0·70	-0·09	0·12	0·02	0·05[$]	0·75[$]	-0·09[$]	0·12[$]
Diploma[a]						1·67	1·49	0·26	-1·25	4·64	2·52	2·07[$]	0·23[$]	-1·52[$]	6·56[$]
Student[b]						7·03*	2·83	0·012	1·63	12·48	6·55*	3·20[$]	0·042[$]	0·36[$]	12·78[$]
Interactions															
SC/IC x Sex											5·75	2·94[$]	0·058[$]	-1·28[$]	11·82[$]
SC/IC x Diploma[a]											-1·56	3·01[$]	0·61[$]	-7·30[$]	4·12[$]
SC/IC x Student[b]											1·29	5·25[$]	0·81[$]	-9·06[$]	11·85[$]
Model fit				df_1	df_2				df_1	df_2				df_1	df_2
R^2	0·01					0·01					0·02				
R^2 adj·	0·01		0·085	1	1,003	0·01		0·057	5	975	0·01		0·055	8	972
Model comparison															
AIC	6,278					6,100					6,101				
BIC	6,288					6,129					6,145				
N	1,005					981					981				

Abbreviations: ß – regression coefficient, SE – standard error· 95% CI – 95% confidence interval· SC – standard care· IC – innovative care.
[a] - Bachelor's/postgraduate degree (ISCED 5-8).
[b] - Currently enrolled in secondary/tertiary/ vocational education.
[$] - based on a bootstrap sample of 4,750.
* - p <0·05.
Model 2, excluding interaction terms, showed the best model fit.

Table S29) Patient satisfaction with care at 12-month follow-up.

Linear regression models statistically comparing standard care and innovative care. Analyses are based on linear regression with bootstrap (repl· 5,000), bias-corrected and accelerated robust standard errors and confidence intervals to counterbalance distortion of the distribution. In addition to the basic model, more complex models, including covariates and interactions, are presented. Analyses were performed using SPSS Version 27.0. Data are presented as correlation coefficients, standard errors and confidence intervals.

	Basic Model					**Covariates Models**									
	Model 1					**Model 2**					**Model 3**				
	ß	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**		**ß**	**SE**	***p* value**	**95% CI**	
SC/IC	1·31*	0·40	0·0014	0·52	2·08	1·35*	0·41	0·0020	0·55	2·15	2·32*	0·73$	0·0016$	0·94$	3·72$
Sex						0·47	0·41	0·26	-0·37	1·28	0·65	0·61$	0·29$	-0·61$	1·92$
Age						-0·02	0·01	0·29	-0·04	0·01	-0·02	0·01$	0·28$	-0·04$	0·01$
Diploma[a]						-0·76	0·46	0·10	-1·68	0·14	0·24	0·68$	0·74$	-1·08$	1·53$
Student[b]						-0·85	0·79	0·27	-2·41	0·66	-0·12	1·01$	0·90$	-2·13$	1·95$
Interactions															
SC/IC x Sex											-0·37	0·82$	0·66$	-1·90$	1·18$
SC/IC x Diploma[a]											-1·86*	0·92$	0·043$	-3·69$	-0·05$
SC/IC x Student[b]											-1·54	1·46$	0·29$	-4·33$	1·29$
Model fit				df_1	df_2				df_1	df_2				df_1	df_2
R^2	0·01					0·02					0·02				
R^2 adj·	0·01*		0·0010	1	945	0·01*		0·0061	5	917	0·01*		0·0071	8	914
Model comparison															
AIC	3,434					3,344					3,346				
BIC	3,444					3,374					3,389				
N	947					923					923				

Abbreviations: ß – regression coefficient· SE – standard error· 95% CI – 95% confidence interval· SC – standard care· IC – innovative care.

[a] - Bachelor's/postgraduate degree (ISCED 5-8).

[b] - Currently enrolled in secondary/tertiary/ vocational education.

$ - based on a bootstrap sample of 3,850.

* - p <0·05.

Model 2, excluding interaction terms, showed the best model fit.