Physica A 392 (2013) 889–901

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

An exploratory statistical approach to depression pattern identification

Qing Yi Feng [a], Frances Griffiths [b,*], Nick Parsons [b], Jane Gunn [c]

[a] Complexity Science Doctoral Training Centre, University of Warwick, Coventry, United Kingdom
[b] Warwick Medical School, University of Warwick, Coventry, United Kingdom
[c] General Practice and Primary Health Care Academic Centre, The University of Melbourne, Melbourne, Australia

ARTICLE INFO

Article history:
Received 16 February 2012
Received in revised form 14 August 2012
Available online 5 November 2012

Keywords:
Depression
Pattern identification
Generalized Information Entropy
Maximum entropy estimates
Quantitative analysis
Visual representation

ABSTRACT

Depression is a complex phenomenon thought to be due to the interaction of biological, psychological and social factors. Currently depression assessment uses self-reported depressive symptoms but this is limited in the degree to which it can characterise the different expressions of depression emerging from the complex causal pathways that are thought to underlie depression. In this study, we aimed to represent the different patterns of depression with pattern values unique to each individual, where each value combines all the available information about an individual's depression. We considered the depressed individual as a subsystem of an open complex system, proposed Generalized Information Entropy (GIE) to represent the general characteristics of information entropy of the system, and then implemented Maximum Entropy Estimates to derive equations for depression patterns. We also introduced a numerical simulation method to process the depression related data obtained by the Diamond Cohort Study which has been underway in Australia since 2005 involving 789 people. Unlike traditional assessment, we obtained a unique value for each depressed individual which gives an overall assessment of the depression pattern. Our work provides a novel way to visualise and quantitatively measure the depression pattern of the depressed individual which could be used for pattern categorisation. This may have potential for tailoring health interventions to depressed individuals to maximize health benefit.

1. Introduction

Depression is a common emotional disorder [1]. For example, in the U.S. depressive disorders affect approximately 18.8 million adults, or about 9.5% of the population age 18 and older, in a given year [2]. According to Australian Government statistics, everyone will at some time in their life be affected by depression — their own or someone else's [3]. The WHO report on mental illness released in 2001 suggests that depression will be the second largest cause of disability burden after heart disease by 2020 [1].

Depression remains a problem for primary care practice and research [4]. Studies over the past decade have consistently found low rates of detection in primary care settings, where most depressed individuals go if they seek care [5]. Until recently there has been little research evidence for how to identify those at risk of chronicity and relapse [6]. However,

* Correspondence to: Warwick Medical School, University of Warwick, Coventry CV4 7AL, United Kingdom. Tel.: +44 24 76522534; fax: +44 24 76528375.
E-mail address: F.E.Griffiths@warwick.ac.uk (F. Griffiths).

0378-4371/$ – see front matter
doi:10.1016/j.physa.2012.10.025

Nomenclature

$\mathbf{x}$	information set, each set contains n information
$d\mathbf{x}$	unit of the ensemble
J	information flux
$\bar{J}$	averaged information flux
t	time
f	constraint
$S_{\bar{J}}$	information entropy
c	integral constant in Eqs. (6)–(8)
x	a piece of information

Greek symbols

ρ	probability density
α	constant in Eqs. (5)–(8)
β	coefficient in Eqs. (5)–(8)
η	constant in Eq. (8)
γ	coefficient in Eq. (8)
σ	coefficient in Eq. (10)
μ	constant in Eq. (10)
Φ	potential function
λ	damping coefficient
ξ	order parameter
ζ	constant in Eq. (13)

Subscripts

i, j, k, l	number to indicate the depressed individual
n	total number of information of the depressed individual

a number of cohort studies [7–9] are now identifying such risk factors in order to develop evidence on how to tailor medical interventions [10], with the aim of preventing chronicity and relapse.

In this paper we describe an exploratory study using an analysis technique from complexity science that has potential to characterise patients individually and so eventually contribute to identifying which treatments are most likely to benefit which patients. We conceptualise depression as a pattern of contributing characteristic factors for each individual at each time point for which data is available. In this exploratory paper we describe analysis for a small number of cases of people with depression. The next step would be to apply this analysis approach to a larger number of cases and link the pattern of depression for each individual with appropriate outcomes. This may then pave the way to undertaking research to evaluate interventions for depression where there is close matching and tailoring of interventions to each individual's depression pattern.

Complexity sciences has provided prospects of achieving the above aim [11], since depressed individuals can be considered as complex systems adjusting, adapting and sometimes transforming with their environments [10]. Patterns of a given system can be investigated by various simulation methods using nonlinear interactions among elements, such as self-organized criticality (SOC), cellular automata (CA), agent-based models, swarm systems, etc. [12–14]. These methods provide an understanding of the emergence of patterns by assuming rules of interaction. For some simple or well-studied systems the assumptions are reasonable, but if we take depressed individuals as systems, it is much more difficult to set up regular rules: as with all people, those living with depression are constantly interacting with their context and interacting with other people. They are also changing with their context over time — the things they have done or experienced influences how they are themselves and their relationships with other people. Thus the pattern of the depressed individual emerges through a dynamic play of interactions including feedback within the body as well as relationships with the physical and social environment [15,16].

Given the difficulties of rule-based simulations, we examined data from depressed individuals, which included assessments of their experiences, wellbeing and context that are known to be strongly related to depression and which may have a causative relationship in themselves or in combination. Using this data we aimed to identify patterns of depression based on the data about each depressed individual. Assessments of depression currently in use mostly ask about depression symptoms alone. Depressive symptoms are an important aspect of patterns of depression and treatment of symptoms can be beneficial. However our aim was to describe patterns of depression based on a wider selection of associated factors. In our study, since we are dealing with information rather than assumed rules, we used statistical analysis to develop a new approach to depression pattern identification.

2. Methods

2.1. Statistical modelling for depression pattern

From the perspective of complexity science, the individual is considered as a "case" in the sociological rather than the medical sense, with health and illness as just one aspect of the individual [10]. Traditional epidemiology and clinical trials use statistical approaches where there is typically an underlying assumption that the individual cases, in our example, the people with depressive symptoms, are independent of one another [17]. This is a reasonable assumption to make for a great deal of health related research even though it could be argued, for example, that there cannot be complete independence of individuals if they live in the same locality or share the same cultural values.

We propose that depressed individuals can be considered as complex systems. Since we cannot isolate depressed individuals from their sociocultural context, we can consider the individuals as subsystems of a larger system, capture all the interactions, and then apply statistical analysis to integrate their information. This analysis approach assumes that individuals living within a locality are not completely independent of each other.

What might the system in this study be like? It is not a closed system described by the law of thermodynamics, in which there is no order or pattern existing when information entropy goes to the maximum. Classic dissipative systems using the concept of negentropy, first introduced to conservative systems by Schrödinger to derive ordered patterns [18,19], still separate the research objects from their environments by boundaries and utilize two information entropy terms to describe the system. This makes it difficult to identify patterns in systems that have more complex interactions. Therefore, the system we need must be an open complex system connected to its environment by information flux which should reflect both internal interactions and external connections of the system.

Suppose we have collected information related to depression from several people, and these people form, in some sense, an open complex system. In the Diamond Cohort Study which we present here, the participants were recruited from 30 general practices in rural and urban areas of the State of Victoria, Australia. Although the individual cohort members are unlikely to interact directly with each other, we suggest they can be assumed to be part of the same complex system as they experience a similar geophysical environment, share similar social structures/organisations, economy and culture, have access to the same broadcast media and access the same health care system. The complex system can be specified if these people had direct interactions with each other but the system can also be generalized if there exist only indirect interactions among them through living within the same State of Australia. The observed data of the depressed individual is represented by the information set $\mathbf{x}_k(x_{k1}, x_{k2}, \ldots, x_{kn})$, where the subscript k indicates the depressed individual.

As in classical statistical mechanics [20], we consider that all possible information microstates of the system comprise of a continuous range — an ensemble, and then $\mathrm{d}\mathbf{x}_k = \mathrm{d}x_{k1}\mathrm{d}x_{k2}\ldots\mathrm{d}x_{kn}$ is a unit of the ensemble. The probability for the state of all units $\mathrm{d}\mathbf{x}$ at time t is $\rho(\mathbf{x}, t)\mathrm{d}\mathbf{x}$, here $\rho(\mathbf{x}, t)$ is the distribution function of the ensemble, which satisfies the normalization condition $\int \rho(\mathbf{x}, t)\mathrm{d}\mathbf{x} = 1$ The parameter J combines all possible information microstates [21–23]:

$$J = \eta + \sum_i \gamma_i \mathbf{x}_i + \sum_{ij} \gamma_{ij}\mathbf{x}_i\mathbf{x}_j + \sum_{ijk} \gamma_{ijk}\mathbf{x}_i\mathbf{x}_j\mathbf{x}_k + \sum_{ijkl} \gamma_{ijkl}\mathbf{x}_i\mathbf{x}_j\mathbf{x}_k\mathbf{x}_l + \cdots. \tag{1}$$

Here η is a constant and γ are coefficients which vary with different interactions within or among $\mathbf{x}_k$. J is the information flux existing within the unit $\mathrm{d}\mathbf{x}$ at time t, the averaged flux function over all possible information microstates is then:

$$\bar{J} = \int \rho\,(\mathbf{x}, t)\,J\,(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{2}$$

Here $\bar{J}$ represents the general characteristics of information entropy of the system and we call it the Generalized Information Entropy (GIE).

Without boundaries, to fully define a system we need constraints. From the perspective of GIE, constraints $\mathbf{f}$ including all kinds of conservations of the system can be generated by certain interactions among information sets $\mathbf{x}$ which can be sufficiently transformed into one to four order momentum quadratures:

$$\langle \mathbf{x}_i \rangle = f_1, \qquad \langle \mathbf{x}_i\mathbf{x}_j \rangle = f_2, \qquad \langle \mathbf{x}_i\mathbf{x}_j\mathbf{x}_k \rangle = f_3, \qquad \langle \mathbf{x}_i\mathbf{x}_j\mathbf{x}_k\mathbf{x}_l \rangle = f_4. \tag{3}$$

Normally constraints are expressed by many partial differential equations (PDEs) with temporal–spatial boundary conditions. However for our system, the constraints are usually unmeasurable, for example people share the same culture; in this scenario, how do we formulate the PDEs? Even if the system has measurable constraints, suppose these people are neighbours and they have regular interactions with each other, given that the interactions are complex, then the PDEs will be mainly nonlinear making numerical solutions difficult to find [23,24]. The treatments for constraints we use are also different from the popular simulation methods mentioned above which set interaction rules to correspond with the non-linear PDEs. Here we believe the constraints can be investigated from the information sets $\mathbf{x}$ — if these people share the same culture or live in the same community they must have something in common, which is reflected by their information set interactions.

Given defined information entropy $\bar{J}$ and constraints, we can implement Maximum Entropy Estimates, or in other words, use the Maximum Entropy Principle to derive equations for depression patterns, by this means depression patterns emerge when the GIE of the whole system goes to a maximum:

$$\begin{cases} S_{\bar{J}} = \max \bar{J} = \int \rho\left(\mathbf{x}, t\right) J\left(\mathbf{x}\right) \mathrm{d}\mathbf{x} \\ \langle \mathbf{x}_i \rangle = f_1, \qquad \left\langle \mathbf{x}_i \mathbf{x}_j \right\rangle = f_2, \qquad \left\langle \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \right\rangle = f_3, \qquad \left\langle \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_l \right\rangle = f_4. \end{cases} \tag{4}$$

Eq. (4) is actually a self-organized optimization model, but unlike other rule-setting models (e.g. Generic Algorithm, Ant Colony Optimization) the optimization here is based on the input information.

By using a Lagrange multiplier Eq. (4) yields:

$$J(\rho) + \rho J'(\rho) - \alpha - \sum_i \beta_i \mathbf{x}_i - \sum_{ij} \beta_{ij} \mathbf{x}_i \mathbf{x}_j - \sum_{ijk} \beta_{ijk} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k - \sum_{ijkl} \beta_{ijkl} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_l - \cdots = 0. \tag{5}$$

Here α is a constant and $\boldsymbol{\beta}$ are coefficients which also reflect the interactions. Then:

$$J = \alpha + \sum_i \beta_i \mathbf{x}_i + \sum_{ij} \beta_{ij} \mathbf{x}_i \mathbf{x}_j + \sum_{ijk} \beta_{ijk} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k + \sum_{ijkl} \beta_{ijkl} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_l + \cdots + \frac{c}{\rho}. \tag{6}$$

Here c is an integral constant. Rearranging Eq. (6) gives:

$$\rho = \frac{c}{J - \alpha - \sum\limits_i \beta_i \mathbf{x}_i - \sum\limits_{ij} \beta_{ij} \mathbf{x}_i \mathbf{x}_j - \sum\limits_{ijk} \beta_{ijk} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k - \sum\limits_{ijkl} \beta_{ijkl} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_l - \cdots}. \tag{7}$$

Substituting Eq. (1) for J in Eq. (7):

$$\rho = \frac{c}{\eta - \alpha + \sum\limits_i \left(\gamma_i - \beta_i\right) \mathbf{x}_i + \sum\limits_{ij} \left(\gamma_{ij} - \beta_{ij}\right) \mathbf{x}_i \mathbf{x}_j + \sum\limits_{ijk} \left(\gamma_{ijk} - \beta_{ijk}\right) \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k + \sum\limits_{ijkl} \left(\gamma_{ijkl} - \beta_{ijkl}\right) \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_l + \cdots}. \tag{8}$$

We obtained the ensemble distribution function of the system. Using the power series for e^x:

$$e^x = 1 + x + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \cdots. \tag{9}$$

We further obtained that the potential function of the system is as follows:

$$\Phi(\sigma, \mathbf{x}) = \mu + \sum_i \sigma_i \mathbf{x}_i + \sum_{ij} \sigma_{ij} \mathbf{x}_i \mathbf{x}_j + \sum_{ijk} \sigma_{ijk} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k + \sum_{ijkl} \sigma_{ijkl} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_l + \cdots. \tag{10}$$

The potential function regulates the asymptotic stabilization of the system. Since the system here is formed by the individuals with certain depressive symptoms, Eq. (10) controls the characteristics of the depression patterns of depressed individuals.

According to the connection of potential functions and dynamic evolutional equations, we deduced the stochastic differential equations (SDEs) of depression pattern from Eq. (10) as:

$$\partial \mathbf{x}_k / \partial t = \lambda_k \mathbf{x}_k + S_k(x_1, x_2, \ldots, x_n) + F_k(t). \tag{11}$$

Here the information entropy term $S_k(x_1, x_2, \ldots, x_n)$ is from the non-linear interactions among information. If the interactions are linear which means all information is independent, there will be no information entropy term and then the equation is exactly the Langevin equation. Therefore, the SDEs deduced by our proposed modelling methodology are the generalized Langevin equations, which are a modified form of the classical Langevin equations. They represent the dynamical evolution of patterns in a system, and the stochastic force $F(t)$, which also appears in the Langevin equations, often makes the pattern jump from one state to another [25].

To derive a potential function related to patterns, we mortify Eq. (10) and diagonalize the constant in second order:

$$\xi_k = \sum_{i=1}^{n} a_{ki} x_{ki}. \tag{12}$$

Here $\boldsymbol{\alpha}$ are coefficients directly related the information x_{ki} of the depressed individual and the logic chain between $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ is: $\boldsymbol{\gamma} \Rightarrow \boldsymbol{\alpha}$. Since $\boldsymbol{\gamma}$ represents the possible information microstates and $\boldsymbol{\alpha}$ information interactions, the logic chain we obtain through statistical modelling sets up a bridge connecting microscopic dynamics and macroscopic characteristics which ensure a more reliable way to identify patterns from observable information rather than assumed rules. And the potential function then becomes:

$$\Phi(\lambda, \boldsymbol{\xi}) = \zeta + \sum_k \lambda_k \xi_k^2 + \cdots. \tag{13}$$

As the main parameter in the potential function, ξ_k is an order parameter defined in Synergetics by Haken [26] which represents the emerged patterns of the system. From the transformation equation (12), it is clear that ξ_k is a combination of an individual's information related to depression. The value of ξ_k can therefore be interpreted as a depression pattern value that combines within it the information that relates to depression for the individual, and the corresponding a_{ki} are the weights or the contribution of information x_{ki} to his/her depression pattern. Thus, with n different kinds of information related to depression in the observed data for each depressed individual in the complex system, individual k has an information set $\mathbf{x}_k(x_{k1}, x_{k2}, \ldots, x_{kn})$ which determines the characteristics of the individual's depression pattern; since we know the weight a_{ki} of information variable x_{ki} by maximizing the GIE of the whole system, we can then characterize the individual's depression by the combination of information variable x_{ki} — the depression pattern value ξ_k.

The potential function equation (13) represents a physical picture of how the system reaches optimization when we implement maximum entropy estimation based on GIE: when the system is in the process of information microstates interacting or allocating information flux, two kinds of patterns will exist: (1) inactive patterns ($\lambda_k < 0$) which cannot obtain enough information flux in the line of least resistance and will be eliminated in the optimization; (2) active patterns ($\lambda_k > 0$) which can obtain sufficient information flux and make the whole system trend towards the optimization state and will survive and emerge.

2.2. Simulation algorithms

Self-Organized Feature Maps (SOMs), a kind of artificial neural network (ANN), are an unsupervised learning mechanism dependent on non-linear interactions among neurons competing for activation and a "winner-take-all" strategy: only the neuron with the highest activation stays active while all other neurons shut down, with other variations allowing more than one neuron to be active also existing. The computational principles of SOMs and the maximum entropy estimate based on GIE are completely equivalent, introducing non-linear connections and optimization to the complex system. Therefore a SOM is an ideal simulation algorithm to analyse depression patterns quantitatively.

In detail, we take the pre-treated information set $\mathbf{x}_k(x_{k1}, x_{k2}, \ldots, x_{kn})$ of the depressed individual k as input to a SOM network. At the beginning, the SOM network has a group of arbitrary numbers x_i^r (or random connection weights w_{ij} among artificial neurons, where w_{ij} is equivalent to the connection among information x_i represented by a_{ki} in Eq. (5)) as the original random state of information x_i. In order to implement the optimization, the SOM network calculates the mean Euclidean distance d between x_i and x_i^r, the smallest distance is the victorious state with corresponding connection weight w_{ij}. SOM networks use positive feedback to adjust x_i^r and repeat this process which is called "training". Training, usually several hundreds of steps when the whole network tends towards stability, results in a final set of connection weights w_{ij} that reflect the contribution of information x_i to the emerged depression pattern of the depressed individual.

2.3. Data source

The data used is the cleaned data from the Diamond Cohort Study, a large longitudinal study approved by The University of Melbourne's Human Research Ethics Committee, conducted in Victoria Australia, which commenced in 2005. Fig. 1 summarises the data collection process: (1) 30 general practices (GPs) in Victoria were recruited from a randomly selected list of 200 GPs provided by the Health Insurance Commission. GPs were eligible to participate if they had seen at least 600 individuals aged 18–75 years in the previous year; were able to generate a computerised list of the details of these people; agreed to complete a survey; and were the only GP in their practice to take part in the study. (2) Practice staff supported by research assistants searched the GP's computerized records to randomly sample 600 eligible individuals from each of the 30 GPs. In total almost 18,000 individuals were sent a screening survey. Individuals were eligible for the screening survey if they: were able to read English; were not terminally ill; were aged 18–75 years; and did not reside in a nursing home. (3) 7667 individuals (43%) returned a completed screening survey. 789 individuals with 'probable clinical depression' who declared an interest in hearing more about the study and had provided their first name and telephone number were recruited into the cohort. (4) Cohort participants were then required to complete a baseline postal survey and follow-up postal surveys were sent at 3, 6, 9 and 12 months. Computer-assisted telephone interviews were undertaken with participants at baseline and 12 months. A full description of data collection is detailed elsewhere [9,27].

2.4. Data selection and pre-treatment

Based on our proposed model, the more exhaustive the information available for the depressed individual, the more accurately the patterns will be identified. However, in practice we cannot collect complete information about a system. For example, a questionnaire cannot cover all aspects of the participant, their life and context. However comprehensive a questionnaire aims to be, participants do not answer all the questions. We suggest that for our objective, to identify patterns of depression and how they change over time, it might only be necessary to have information about each individual that is known to be important in relation to depression. From the variables collected as part of the Diamond Cohort study, we selected 14 factors known to be strongly related to depression [6,7] as the information set $\mathbf{x}$, shown in Table 1. Some variables

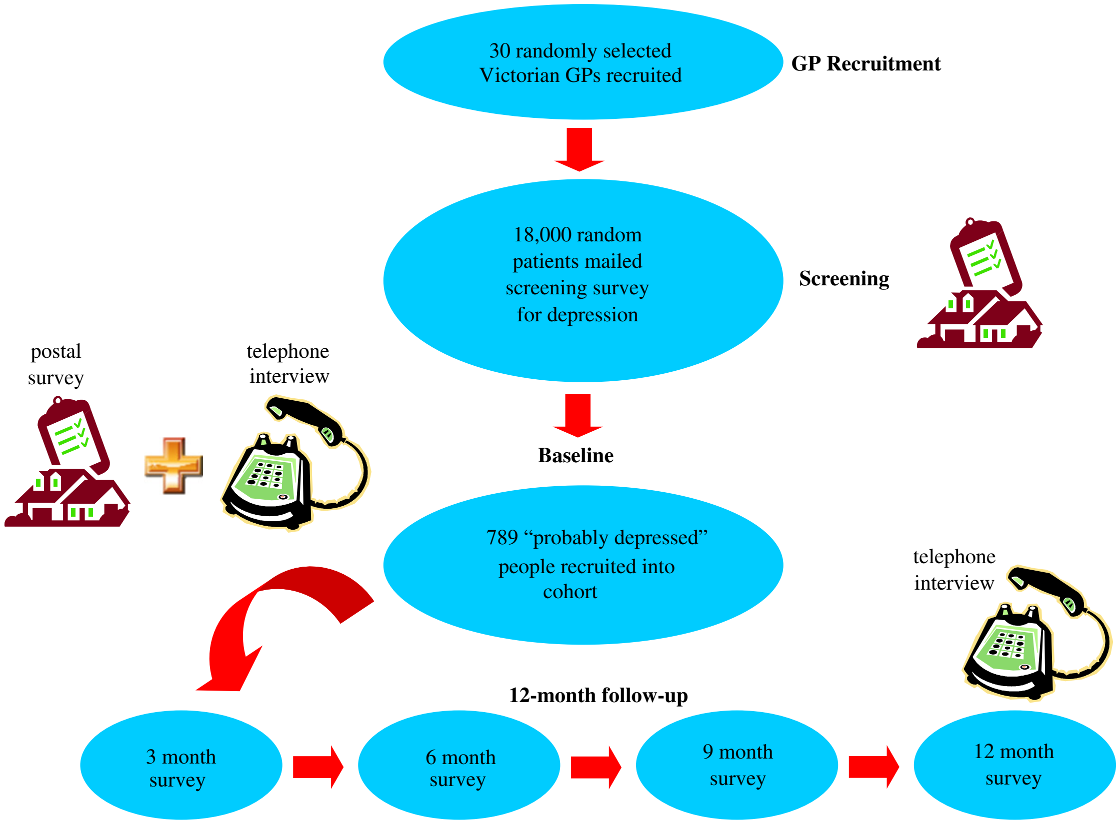

Fig. 1. Flow chart of the collection of depression related data in the Diamond Cohort.

were used as collected from participants, e.g. "employment status", but others were formed by combining a number of variables, e.g. the value of "total exercise" is the sum of "vigorous exercise" and "moderate exercise".

In order to ensure the final depression pattern values consistently quantified the state of depression, we reversed the values of some variables so a higher value of the variable always contributed to increasing the likelihood of depression. For example, in the original data for "social relationship" 100 represented a person with excellent social relationship so this was reversed so that 100 meant this person had no social relationships at all. Also as a requirement for neural network simulation, we rescaled all the data to the interval $[0, 1]$.

3. Analysis

Technically by using the methodology described above we can identify the pattern of every depressed person involved in the Cohort since all the participants form a large complex system. In this paper, we present exploratory analysis of information about 28 individuals from the cohort. These individuals were chosen for this exploratory analysis as based on a mixed methods longitudinal analysis concurrently underway these cases were known to have a diversity of depression trajectories over one year.

3.1. Numerical simulation of different systems

By utilizing the baseline data of the chosen 28 depressed individuals we formed a small complex system to identify their depression patterns. Since we chose 14 variables according to Eq. (5) and normalised them into the range of $[0, 1]$, the value of ξ_k in this study ranges from 0 to 14 with a higher value indicating a more severe state of depression. We then implemented our proposed simulation algorithms using SOMs by choosing a network of 4×7 and 500 training steps.

After numerical simulation, we obtained an evolutionary figure to show how ξ_k values change in the SOM training process and a set of connection weights w_{ij} of variables for the depressed individual. We averaged the last 100 stable w_{ij} values as the contribution of information x_i to the emerged depression pattern represented by a_{ki} in Eq. (5). Since every set of connection weights can be used to identify a unique pattern for the depressed individual, we called this identified pattern the "depression fingerprint" of the depressed individual.

Table 1
14 major factors highly related to depression.

Variables number[*]	Variables name	Score range and meaning	Whether reversed
1	Employ status	[1, 3] 1 = "Employed/student" 2 = "Not employed/not in paid employment" 3 = "Unable to work"	No
2	Live alone status	[0, 1] 0 = "No" 1 = "Yes"	No
3[a]	Income management	[1, 5] 1 = "Easily" 2 = "Not to bad" 3 = "Difficult some of the time" 4 = "Difficult all of the time" 5 = "Impossible"	No
3[b]	Money meets needs	[1, 5] 1 = "Not at all" 2 = "Slightly" 3 = "Somewhat" 4 = "To a great extent" 5 = "Completely"	Yes
4	Environmental score	[0, 100] 0 = "very poor environmental status" 100 = "excellent environmental status"	Yes
5	Social relationship	[0, 100] 0 = "no social relationship at all" 100 = "excellent social relationship"	Yes
6	Psychological score	[0, 100] 0 = "very poor psychological status" 100 = "excellent psychological status"	Yes
7	Physical score	[0, 100] 0 = "very poor physical status" 100 = "excellent physical status"	Yes
8[a]	Somatic symptoms	[0, 30] 0 = "no somatic symptoms at all" 30 = "having all the measured somatic symptoms"	Yes
8[b]	Absent days because of physical problems	[0, 273.75] 0 = "no absent day at all" 273.75 = "be absent everyday in a period of 9 months"	No
9	Total exercise	[0, 10] 0 = "no exercise at all" 10 = "doing exercise every day"	Yes
10	Community participation	[0, 90] (Baseline survey) [0, 70] (Follow-up surveys) 0 = "no community participation at all" 90 or 70 = "having community participation very often"	Yes
11	Recent life positive impacts	[0, 13] (Baseline survey) [0, 18] (Follow-up surveys) 0 = "no positive impacts at all" 13 or 18 = "had every measured positive impacts"	Yes
12	Recent life negative impacts	[0, 13] (Baseline survey) [0, 18] (Follow-up surveys) 0 = "no negative impacts at all" 13 or 18 = "had every measured negative impacts"	No
13	Total childhood sexual abuse	[0, 3] 0 = "no childhood sexual abuse history" 3 = "had childhood sexual abuse more than twice"	No
14	Total childhood physical abuse	[0, 6] 0 = "no childhood physical abuse history" 3 = "had childhood physical abuse more than 5 times"	No

[*] These numbers 1–14 are used in Figs. 3 and 4 representing the same variables.
[a] The variable was used in the numerical simulations of temporal changes of depression fingerprint.
[b] The variable was used in the numerical simulations of temporal changes of depression pattern.

Furthermore, from Eq. (11) we know that the shapes of "depression fingerprint" of the depressed individual will change over time. Thus if we consider the depressed individual as one temporal complex system then we can utilize his/her data at different time points to analyse the changes of his/her depression pattern over time. Here for every depressed person we had 5 time points: baseline, 3 month, 6 month, 9 month and 12 month.

Table 2
Depression pattern values of the depressed individual at baseline and standard deviations (SD) with CES-D score in the ascending order.

Study id	Depression pattern value (range 0–14)	SD	CES-D score (range 0–60)
9 026	2.16	0.06	1
8 155	1.98	0.05	2
8 274	2.34	0.10	3
16 237	3.41	0.30	5
21 178	2.53	0.07	8
11 055	4.14	0.14	9
30 026	4.04	0.13	11
11 067	4.84	0.12	15
13 056	4.36	0.18	15
23 092	2.98	0.11	15
15 135	2.78	0.09	20
8 096	3.53	0.16	21
6 087	2.94	0.17	23
4 108	2.62	0.12	26
26 041	5.17	0.15	26
11 118	4.09	0.20	27
20 038	4.80	0.12	28
22 074	4.10	0.13	30
6 057	4.71	0.14	31
11 088	5.73	0.24	32
4 020	5.47	0.24	36
8 286	7.46	0.33	39
3 062	4.90	0.13	42
32 125	5.48	0.28	45
18 238	5.98	0.23	47
26 087	3.61	0.16	47
7 171	5.45	0.15	54
8 051	7.05	0.33	54

3.2. *Comparison with traditional methods*

The 20-item Centre for Epidemiologic Studies Depression Scale (CES-D) is a common measurement of depressive symptoms used in the community, and scores range from 0 to 60 with high scores indicating higher depressive symptoms [28]. We averaged the last 100 steady ξ_k values as the depression pattern value of the depressed individual and compared this to his/her CES-D score (see Table 2). The standard deviation (SD) of ξ_k is also provided to show the stabilities of the emerged patterns.

4. Results

4.1. *Depression pattern features in baseline*

It can be seen in Fig. 2 that after several steps the ξ_k values gradually approach an attractor and remain within a relatively narrow range. Since ξ_k is the order parameter in the system, Fig. 2 also presents a precise picture of how the depression pattern of the depressed individual emerges through information interactions and gives a visual view of pattern formation dynamics.

Unlike the CES-D score, the depression pattern value is a unique value for each depressed individual in the system and gives an overall assessment of the pattern. This depression pattern value could be used for pattern categorization. For example, the individuals labelled as study id 7171 and study id 8051 have the same CES-D score 54, suggesting a similar level of depressive symptoms and as such they may be treated medically in the same manner, while their depression pattern values of 5.45 and 7.05 respectively suggests that they probably have very different depression experiences which may require quite different treatment approaches. Conversely, study id 4020 with depression pattern value 5.47 may be classified as similar to the case with study id 7171 (depression pattern value 5.45) although their CES-D scores differ by 18 points.

In Fig. 3 we show four depression fingerprints. The solid areas in the circles indicate dysfunction, so the greater the size of the solid area the more dysfunction, the smaller the solid area the less dysfunction. The fingerprints in Fig. 3(A) and (D) are for individuals with a similar CES-D score (id 4020: 36 and id 8286: 39) yet the depression pattern value is different (5.47 and 7.46 respectively). The difference in pattern is also clear on viewing the fingerprints. Similarly in Fig. 3(B) and (C), the individuals have the same CES-D score (54) but different pattern values (id 7171: 5.45 and id 8051: 7.05). Again the difference in depression pattern is clear on viewing the fingerprint. The shape of the fingerprint represents the factors contributing to the depression that were included in the analysis. In Fig. 3(B) and (C) for example: for study id 7171, social factors (living alone, poor social relationships, lack of community participation), physical factors (poor physical status, lack

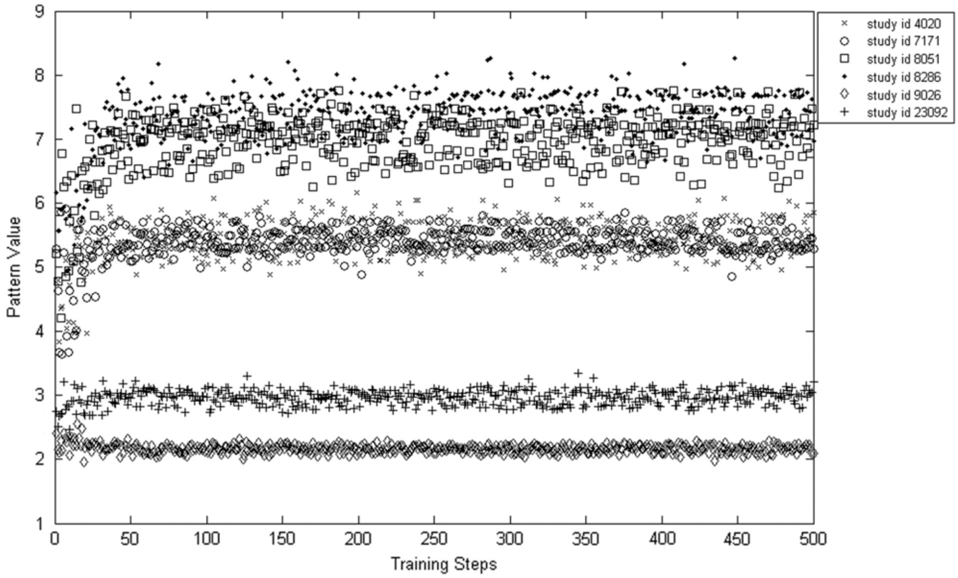

Fig. 2. How ξ_k value changes in the SOM training process.

of exercise) and financial problems (poor income management) are major contributors. For study id 8051 childhood abuse and poor environmental status are prominent.

4.2. Temporal changes of depression patterns

Fig. 4 shows the temporal changes of depression pattern for depressed individuals together with values comparing with corresponding CES-D scores over 5 different time points. The trend of temporal depression value is similar to the one of CES-D score but with less fluctuation.

5. Discussion

5.1. Practical application of depression pattern values

There is ongoing controversy and debate about using symptom-based diagnostic criteria for depression without taking into account how the individual functions day to day, or other factors which may be associated with a poor outcome [17,29]. Clinicians usually diagnose depression based on symptoms often supported by a patient-completed questionnaire about symptoms such as the PHQ-9, HADS or the BDI-II [30–33]. Rather than focusing on symptoms we devised a modelling methodology to draw together multiple types of information related to factors strongly associated with increased risk of depression, and which may have causative influence, including functional measures, perceived health status, social measures and risk factors. We introduced a novel way to identify a depression pattern by calculating depression pattern values ξ_k and weights a_{ki} that represents the importance of a certain factor related to depression. We were also able to demonstrate that the depression pattern values obtained varied over time even though some of the variables as measured are unchanging (e.g. having had the experience of childhood abuse). If the data were available, a measure of the impact of the childhood experience might provide a variable more sensitive to change. It is arguable whether the 14 variables we used here are sufficient to identify depression patterns with utility for clinical practice. However, these variables were selected to represent a biopsychosocial and functional model of depression causation. There is some evidence that the depression pattern values changed over time with a similar trend to the of CES-D scores. As would be expected, the depression value that includes multiple factors, some of which do not change or change slowly, fluctuates differently to symptoms.

5.2. Practical application of "depression fingerprint" and the temporal changes

Assessing patients as individuals is part of the traditional art of doctoring [34], and patients often want advice that is individually tailored [35]. Our analytical approach suggests that it is possible to identify for each depressed person within a particular complex system (for example, the community of the State of Victoria, Australia), a unique "depression fingerprint". The "depression fingerprint" provides both a numerical score and a visual pattern. The fingerprint could potentially be used to tailor interventions to the fingerprint pattern and track change over time for individuals. In research this analysis method may have potential for identifying subgroups of depression.

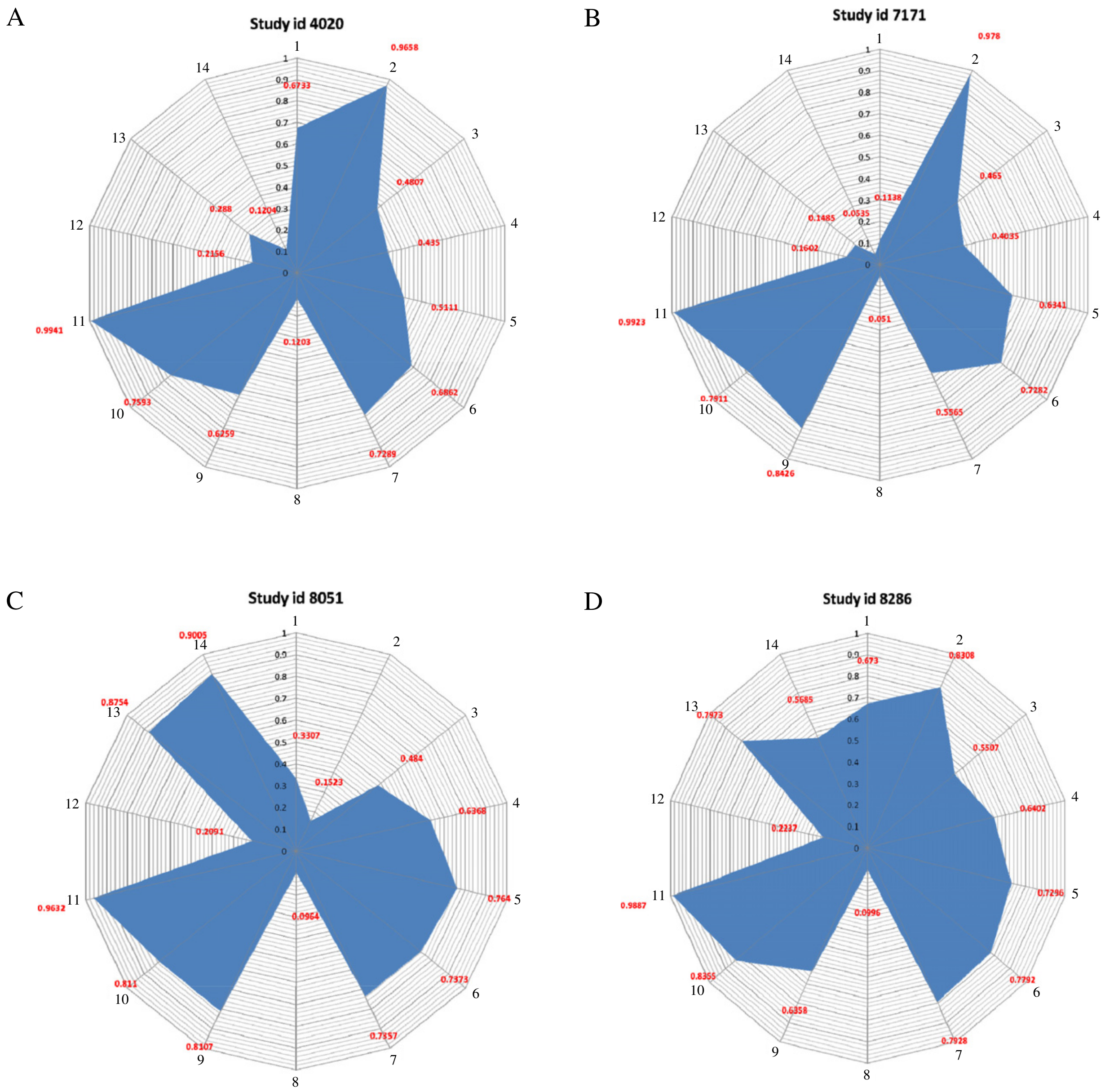

Fig. 3. Four depression fingerprints of depressed individuals (in every fingerprint, the numbers 1–14 stand for the 14 variables in Table 1, numbers (all less than 1) are the stable values of connection weights w_{ij} of variables, and solid area in every fingerprint shows dysfunction the depressed individual). (A) Depression fingerprint for study id 4020. (B) Depression fingerprint for study id 7171. (C) Depression fingerprint for study id 8051. (D) Depression fingerprint for study id 8286.

5.3. Future research

The results presented in this paper are only preliminary; however the methodology proposed here has shown promise and has clinical relevance. Future research is needed to test the potential for this approach to analysis including analysis of information from large numbers of individuals, evaluation of the temporal changes of depression patterns, and relating the depression patterns to depression outcomes.

6. Conclusions

As a common emotional disorder, depression remains a problem for primary care both in practice and research. Considering different interventions might be appropriate for different types of depression, so it is important to develop

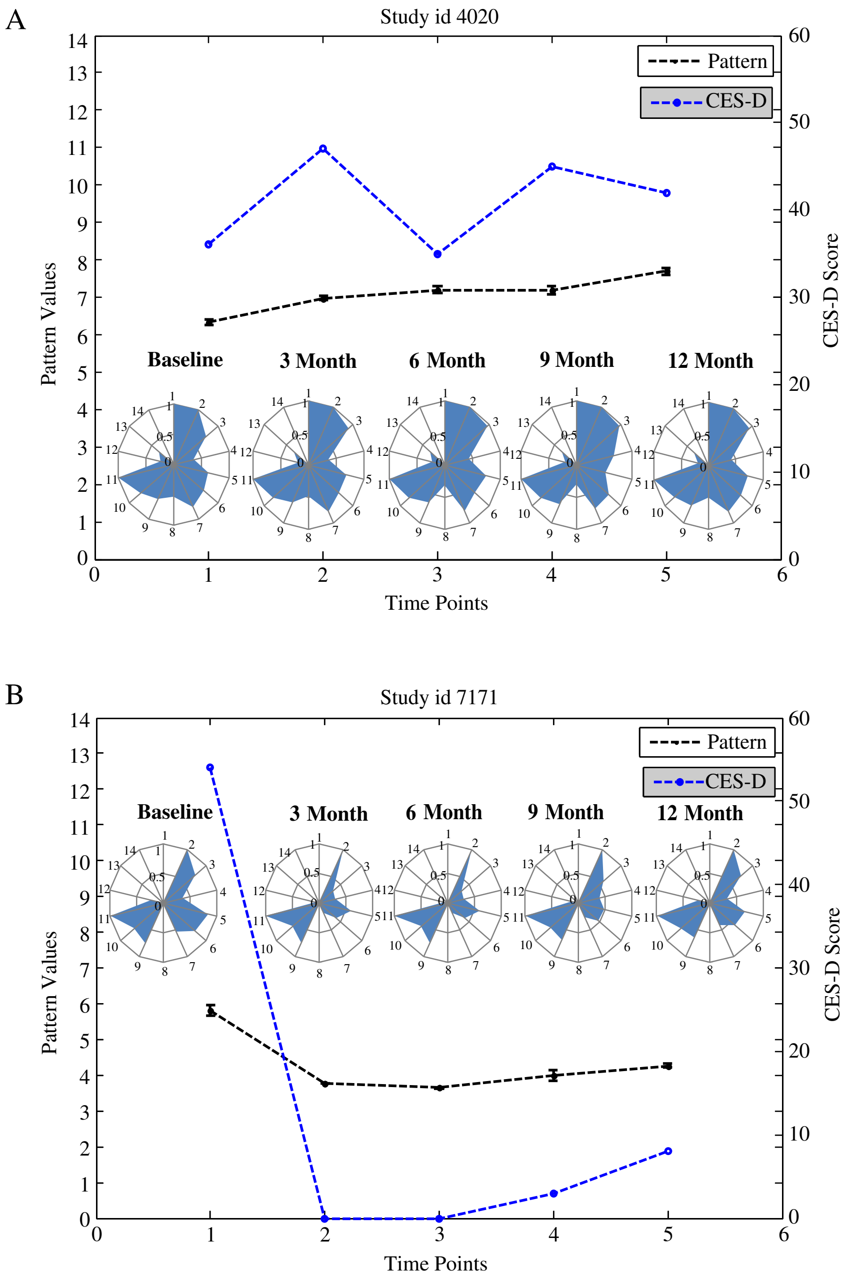

Fig. 4. The temporal changes of depression pattern for depressed individuals together with values comparing with corresponding CES-D scores over 5 different time points (the subplot figures in each figure are the shapes of depression patterns of the depressed person). (A) The temporal changes of depression pattern of study id 4020. (B) The temporal changes of depression pattern of study id 7171. (C) The temporal changes of depression pattern of study id 8051.

evidence on whether an intervention is appropriate to a certain type of depression. In this paper we aimed to identify depression patterns by utilizing the data of depressed individuals about how they functioned and related to their context.

This paper suggests that the depressed individual can be considered as subsystems of an open complex system. This complex system is connected with its environment by its information flux which describes the status of the system both internally and externally to the environment. By averaging the information flux function over all possible information microstates, the Generalized Information Entropy has been defined. By implementing Maximum-Entropy Estimation (used the Maximum Entropy Principle) to derive equations for depression patterns and the stochastic differential equations we deduced what we believed to be a modified form of the classical Langevin equations. Based on the proposed model, we also

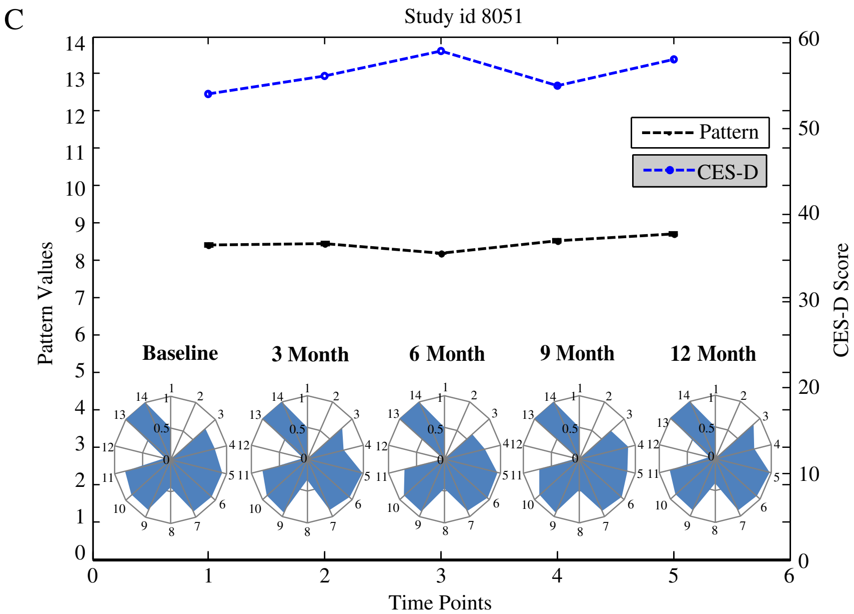

Fig. 4. (*continued*)

developed a numerical simulation method using Self-Organized Feature Maps (SOMs) to process the depression related data obtained by the Diamond Cohort Study.

Our work devised a modelling methodology to integrate the depression information of depressed individuals and introduced a novel way to identify individual depression patterns by calculating depression pattern values and the weights which represent the importance of a certain factor related to depression. We also developed a visual method to show a unique "depression fingerprint" of the depressed individual and temporal changes of depression pattern which may provide practical applications for tailoring medical interventions for individuals with depression.

Acknowledgments

The Diamond study was initiated with pilot funding from the beyondblue Victorian Centre of Excellence and the main cohort has received project grant funding from the National Health and Medical Research Council (ID 299869, 454463, 566511 & 1002908). The one-year Computer Assisted Telephone Interview was funded by a Stream 3 grant from the Australian Primary Health Care Research Institute (APHCRI). No funding body had a role in study design; the collection, analysis, and interpretation of data; the writing of the manuscript; or the decision to submit this manuscript for publication. We acknowledge the 30 dedicated GPs, their patients and practice staff for making this research possible. We thank the cohort participants for their ongoing involvement in the study. We also thank the Diamond project investigators and project team staff who provided assistance in preparing the data for analysis and providing the data files. Furthermore, we thank Dr. Victoria Mason for editing the paper.

References

[1] WHO, The world health report 2001 — Mental health: new understanding, new hope, World Health Organization, Geneva, 2001.
[2] NIMH, Depression research, National Institute of Mental Health, Bethesda, 1999.
[3] AIHW, National health priority areas mental health: a report focusing on depression, Australian Institute of Health and Welfare, Canberra, 1998.
[4] D. Pilgrim, The survival of psychiatric diagnosis, Social Science & Medicine 65 (2007) 536–547.
[5] RAND Corporation, The societal promise of improving care for depression, Rand Corporation, Santa Monica, 2008.
[6] G. Gilchrist, J. Gunn, Observational studies of depression in primary care: what do we know? BMC Family Practice 8 (2007) 1–18.
[7] H. Herrman, D.L. Patrick, P. Diehr, et al., Longitudinal investigation of depression outcomes in primary care in six countries: the LIDO study, functional status, health service use and treatment of people with depressive symptoms, Psychological Medicine 32 (2002) 889–902.
[8] J. Gunn, Depression as a chronic and disabling experience, the DIAMOND study: a longitudinal naturalistic study of depression in primary care in Victoria, Australia, Journal of Affective Disorders 91 (2010) S21–S22.
[9] J.M. Gunn, G.P. Gilchrist, P. Chondros, M. Ramp, K.L. Hegarty, G.A. Blashki, D.C. Pond, M. Kyrios, H.E. Herrman, Who is identified when screening for depression is undertaken in general practice? Baseline findings from the Diagnosis, Management and Outcomes of Depression in Primary Care (diamond) longitudinal study, Medical journal of Australia 188 (12 Suppl) (2008) S119–S125.
[10] F. Griffiths, J. Borkan, D. Byrne, et al., Developing evidence for how to tailor medical interventions for the individual patient, Qualitative Health Research 20 (2010) 1629–1641.
[11] M. Lima, Visual Complexity: Mapping Patterns of Information, Princeton Architectural Press, New York, 2011.
[12] P. Bak, How Nature Works: The Science of Self-Organized Criticality, Copernicus, New York, 1996.
[13] M.I. Rabinovich, A.B. Ezersky, P. Weidman, The Dynamics of Patterns, World Scientific, Singapore, 2000.

[14] J.G. Bazan, J.F. Peters, A. Skowron, Behavioral Pattern Identification Through Rough Set Modelling, in: Lecture Notes in Computer Science, vol. 3642, 2005, pp. 688–697.
[15] C. Battersby, The Phenomenal Woman: Feminist Metaphysics and the Patterns of Identity, Routledge, New York, 1998.
[16] D. Meyers, Feminist perspectives on the self, in: E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, spring 2010 ed., 2004, Available from: http://plato.stanford.edu/archives/spr2010/entries/feminism-self/.
[17] H. Herrman, M. Maj, N. Sartorius, Depressive Disorders, third ed., John Wiley & Sons, Chichester, 2009.
[18] E. Schrodinger, What is Life? With Mind and Matter and Autobiographical Sketches, Cambridge University Press, Cambridge, 1992.
[19] I. Prigogine, From Being to Becoming, W.H. Freeman, San Francisco, 1980.
[20] E.T. Jaynes, Information theory and statistical mechanics, Physical Review 106 (1957) 620–630.
[21] L. Chai, M. Shoji, Self-organization and self-similarity in boiling systems, Journal of Heat Transfer 124 (2002) 507–515.
[22] Q. Feng, L. Chai, A new statistical dynamic analysis on vegetation patterns in land ecosystems, Physica A: Statistical Mechanics and its Applications 387 (2008) 3583–3593.
[23] M.C. Cross, P. Hohenberg, Pattern formation outside of equilibrium, Reviews of Modern Physics 65 (1993) 851–1112.
[24] S.A. Kauffman, Investigation, Oxford University Press, New York, 2000.
[25] L. Chen, L. Chai, A theoretical analysis on self-organized formation of microbial biofilms, Physica A: Statistical Mechanics and its Applications 370 (2006) 793–807.
[26] H. Haken, Information and Self-Organization, Springer, Berlin, 2000, second enlarged.
[27] J.M. Gunn, D.R. Ayton, K. Densley, et al., The association between chronic illness, multimorbidity and depressive symptoms in an Australian primary care cohort, Social Psychiatry and Psychiatric Epidemiology (2010) http://dx.doi.org/10.1007/s00127-010-0330-z. (Epub ahead of print).
[28] L.S. Radloff, The CES-D scale: a self-report depression scale for research in the general population, Applied Psychological Measurement 1 (1977) 385–401.
[29] C. Dowrick, Beyond Depression: A New Approach to Understanding and Management, Oxford University Press, Oxford, 2004.
[30] K. Kroenke, R.L. Spitzer, J.B. Williams, The PHQ-9: validity of a brief depression severity measure, Journal of General Internal Medicine 16 (2001) 606–613.
[31] A.S. Zigmond, R.P. Snaith, The hospital anxiety and depression scale, Acta Psychiatrica Scandinavica 67 (1983) 361–370.
[32] A.T. Beck, R.A. Steer, R. Ball, et al., Comparison of Beck depression inventories-IA and -II in psychiatric outpatients, Journal of Personal 67 (1996) 588–597.
[33] National Institute for Health and Clinical Excellence, Common mental health disorders: identification and pathways to care. CG123. London: National Institute for Health and Clinical Excellence, 2011. http://publications.nice.org.uk/common-mental-health-disorders-cg123/guidance#step-1-identification-and-assessment.
[34] E. Cassell, The Nature of Suffering and the Goals of Medicine, Oxford University Press, Oxford, 2004.
[35] R. Massé, F. Légaré, The limitations of a negotiation model for perimenopausal women, Sociology of Health and Illness 23 (2001) 44–64.