GeroScience (2017) 39:551–556
DOI 10.1007/s11357-017-9996-4

ORIGINAL ARTICLE

The application of information theory for the estimation of old-age multimorbidity

David Blokh · Ilia Stambler · Emilia Lubart · Eliyahu H. Mizrahi

Received: 16 August 2017 / Accepted: 23 August 2017 / Published online: 28 August 2017

Abstract Elderly patients are commonly characterized by the presence of several chronic aging-related diseases at once, or old-age "multimorbidity," with critical implications for diagnosis and therapy. However, at the present there is no agreed or formal method to diagnose or even define "multimorbidity." There is also no formal quantitative method to evaluate the effects of individual or combined diagnostic parameters and therapeutic interventions on multimorbidity. The present work outlines a methodology to provide such a measurement and definition, using information theoretical measure of normalized mutual information. A cohort of geriatric patients, suffering from several age-related diseases (multimorbidity), including ischemic heart disease, COPD, and dementia, were evaluated by a variety of diagnostic parameters, including static as well as dynamic biochemical, functional-behavioral, immunological, and hematological parameters. Multimorbidity was formally coded and measured as a composite of several chronic age-related diseases. The normalized mutual information allowed establishing the exact informative value of particular parameters and their combinations about the multimorbidity value. With the currently intensifying attempts to reduce aging-related multimorbidity by therapeutic interventions into its underlying aging processes, the proposed method may outline a valuable direction toward the formal indication and evidence-based evaluation of effectiveness of such interventions.

Keywords Multimorbidity · Frailty · Aging · Aging-related diseases · Normalized mutual information

D. Blokh
C.D. Technologies Ltd., Beersheba, Israel

I. Stambler (✉)
Department of Science, Technology and Society, Bar Ilan University, 5290002 Ramat Gan, Israel
e-mail: ilia.stambler@gmail.com

E. Lubart · E. H. Mizrahi
Geriatric Medicine and Rehabilitation Department B, The Geriatric Medical Center "Shmuel Harofe", Beer Yaakov, Affiliated to Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel

Introduction

With the growing aging population, and the accompanying growing incidence of aging-related diseases, there is an increasing need to improve the diagnostic and therapeutic capabilities for those diseases (Jin et al. 2015). However, the diagnosis and treatment of elderly patients is often greatly complicated by the presence in these patients of several chronic aging-related diseases at once, or age-related "multimorbidity." Currently, there is no agreed or formal method to diagnose or even

define "multimorbidity" (Salive 2013). There are also no agreed and formal methods to evaluate the diagnostic ability of particular diagnostic parameters or their combinations to diagnose multimorbidity. There is an urgent need to develop such quantifiable definitions of multimorbidity and of the effects of diagnostic and therapeutic factors for the elderly population. This need is emphasized by the growing realization of the determinative impact of degenerative aging processes on the emergence of age-related multimorbidity (Rae et al. 2010; Fontana et al. 2014; Goldman et al. 2013). There are now intensifying attempts to reduce old-age multimorbidity by therapeutically intervening into its underlying aging processes (Hall 2015; Newman et al. 2016). However, in order to estimate the success or failure of such interventions, there is a critical need to be able to reliably and quantitatively evaluate the multimorbidity and dynamic changes in it. The present work outlines a methodology to provide such a capability, using information-theoretical measure of normalized mutual information.

It is yet far from solving the problem of precise quantitative diagnostic evaluation of old-age multimorbidity. However, the information-theoretical methodology may offer some directions for the solution. This is due to the fact that information-theoretical measures (such as normalized mutual information, as employed in this study) allow the researchers to evaluate the exact quantitative correlation between any combined group of parameters (such as a combined group of diagnostic markers) with any other combined group of parameters (such as multimorbidity composed from several diseases). Moreover, the information-theoretical methodology uniquely permits the evaluation of cumulative, synergistic, or holistic relations between such combinations (Blokh and Stambler 2016). Such synergistic effects of combinations of parameters are impossible to establish by other methodologies, such as simple arithmetic "scoring" that is sometimes used for the evaluation of multimorbidity.

This capability is exemplified here using a data base on geriatric patients, including some routinely available biochemical, cellular and physiological parameters, and several prevalent age-related diseases, namely ischemic heart disease (IHD), dementia, and COPD. Multimorbidity is here formally coded and measured as a composite of several chronic age-related diseases, in fact producing a new single disease entity—"the multimorbidity"—out of several diseases (see the section Mathematical analysis below).

Methods

Mathematical analysis

In this work, in order to formally measure old-age multimorbidity, we use the information-theoretical measure of correlation of individual or combined diagnostic parameters with individual diseases or multimorbidity, namely the measure of normalized mutual information (NMI), also known as the uncertainty coefficient. The normalized mutual information value tells the exact amount of information that each diagnostic parameter or combination of parameters contain about the multimorbidity value.

Briefly, the normalized mutual information is determined as follows. Let X be a discrete random value with a distribution function.

X	x_1	x_2	.....	x_n
Q	p_1	p_2	.....	p_n

X—can be a biomarker, n—the number of categories of the marker, p_i—the frequency of the category x_i. The entropy of random value X is:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i$$

Let X,Y be the discrete random values (parameters). The algorithms for the determination of normalized mutual information between parameters or their combinations have been presented earlier (Blokh and Stambler 2015a; Blokh and Stambler 2017). Very briefly, for the parameters X,Y, we calculate the value of normalized mutual information c, by the following formula:

$$c = \frac{I(X;Y)}{H(Y)} = \frac{H(X) + H(Y) - H(X,Y)}{H(Y)}$$

where $H(X)$, $H(Y)$, and $H(X,Y)$ are the entropies of random variables X, Y, and $X \times Y$, respectively. The values of the uncertainty coefficient (normalized mutual information) closer to zero indicate a smaller degree of correlation, while the coefficient values closer to 1 indicate a larger degree of correlation.

In order to estimate the correlation between a combined marker and a single disease, we need to estimate the combined correlation of all the markers comprising the combined marker with the disease. For a combined

marker comprised of two markers, this is done in the following way: Let the combined marker Z be comprised of two discrete markers z_1 and z_2, while the marker z_1 assumes two values: 0 and 1, and the marker z_2 assumes three values: 0, 1, and 2. Then, the correlation of the combined marker Z with the disease is estimated by the correlation of a "single marker" assuming 6 values in accordance to the values of the single markers z_1 and z_2: (0,0) – 0, (0,1) – 1, (0,2) – 2, (1,0) – 3, (1,1) – 4, (1,2) – 5. We can proceed in the same way for combined markers comprised by more than two markers.

In this study, instead of a single disease (e.g., IHD), we introduce a new disease entity—the "multimorbidity." The diseased state of "multimorbidity" is a composite of the diseases the patient has. For the present study, the composite "multimorbidity" variable is composed of three diseases: IHD, chronic obstructive pulmonary disease (COPD), and dementia. But the composite "multimorbidity" can be comprised from any number of diseases and morbid or disabled states as relevant for the particular study or clinical setting. We code the "multimorbidity" as a single discrete disease entity assuming several possible states, according to the presence or absence of particular diseases, as follows:

IHD	COPD	Dementia	Multimorbidity
0	0	0	0
1	0	0	1
0	1	0	2
1	1	0	3
0	0	1	4
1	0	1	5
0	1	1	6
1	1	1	7

Here, 0 indicates absence of a disease, and 1 indicates presence of a disease. The final column indicates the state codes for the multimorbidity variable. For example, if both IHD and dementia are present, but not COPD, then the multimorbidity code is 5. In other words, we consider a single composite disease—the "multimorbidity"—that can assume 8 discrete states.

Case materials

This work is based on the analysis of 197 patients (male and female, aged 63–97) treated for hip fracture at the Geriatric Medical Center "Shmuel Harofe" in Beer Yaakov, Israel. Access to the patients' data was obtained according to the principles of the Declaration of Helsinki. Out of all the analytical parameters on the patients, several parameters were chosen to illustrate the methodology presented in this article. The parameters were selected to represent different kinds of analysis: microelements (potassium–K and sodium–Na levels), cellular/immunological (number of lymphocytes–Lym and white blood cells–WBC), hematological (number of thrombocytes–Thr, hemoglobin–Hb), physiological (heart rate–pulse), metabolic (glucose), and functional/behavioral evaluations, namely different types of functional independence measurements–FIM, such as total FIM (Tfim), cognitive FIM (Cfim), and motor FIM (Mfim) (Linacre et al. 1994; Dodds et al. 1993).

Data were evaluated at admission (ad) and discharge from the hospital (dis). Moreover, based on the admission and discharge data, the dynamic change and stability of the parameters were estimated for several parameters as percent positive or negative change above and below a certain threshold: Lym + 10%, Lym − 10%, Thr + 10%, Thr − 10%, Pulse + 5%, Pulse − 5%, Na + 1%, Na − 1%, Gluc + 10%, Gluc − 10%, Hb > + 15%, Hb > − 15%, WBC + 1%, WBC − 1%. Increases or decreases beyond the threshold boundaries may indicate excessive instability of the parameters, potentially indicative of impaired homeostatic/regulatory capacity of the organism. For example, the parameter Lym + 10% was assumed to equal 1, if during the hospital stay, the amount of lymphocytes increased by 10% and more, and 0 otherwise. The parameter Lym − 10% was assumed to be 1, if during the hospital stay, the number of lymphocytes decreased by 10% and more, and 0 otherwise. The thresholds were selected according to the algorithm for boundaries determination by maximizing normalized mutual information (Blokh and Stambler 2015b) for the entire patients' cohort. The patients' age and gender were also included as necessary discriminative parameters. Altogether, 42 parameters are considered. The parameters and the parameters' combinations were correlated with the multimorbidity variable, composed of three degenerative diseases of different organ systems: IHD, COPD, and dementia. The multimorbidity variable was coded as a single disease that can assume eight states, according to the presence or absence of each of the particular diseases.

Results and discussion

Using the above methodology, first we correlated individual diagnostic parameters with the multimorbidity variable (assuming eight states according to the presence or absence of particular diseases—IHD, COPD, and dementia. The results are shown in Table 1. As it can be seen in Table 1, the most informative parameters for the evaluation of multimorbidity were the functional evaluations—cognitive functional independence measurement (Cfim) at admission and at discharge, total functional independence measurements (Tfim) at admission and discharge, and motor functional independence measurements (Mfim) at discharge. These functional parameters have the highest informative values (NMI), ranging from 0.0954 for Cfimdis down to 0.05406 for Tfimad.

Generally, the biomarkers—including the cellular, immunological, biochemical, and physiological parameters—were found to be less informative than the behavioral-functional measurements. Among the biomarkers, following the above functional measurements, the highest NMI value was found for glucose at admission and discharge – 0.05372 and 0.0441, respectively. This may emphasize the important role of glycation as a fundamental mechanism of aging for the emergence of multiple aging-related diseases (Semba et al. 2010). Incidentally, the recently quite famous TAME study—"Targeting Aging with Metformin" that aims to reduce age-related multimorbidity by intervening into its underlying aging process, utilizes the well-known antidiabetic biguanide drug—metformin, which is a "glucophage" with a primary function of reducing glucose and the corresponding glycation (Hall 2015; Newman et al. 2016). The present findings further emphasize the potential role of glucose levels as an indicator or predictor of multimorbidity.

Still, the functional parameters were more informative. This may be partly explained by the specific nature of the functional tests for the diseases under consideration (e.g., the specific implications of cognitive functional independence measurements for the presence of dementia). Being relatively inexpensive and easily applicable by the geriatric physicians, the functional tests can thus provide a good and convenient indication of the multimorbidity status. The use of the biomarkers was more uncertain as indicators of the multimorbidity. Among the biomarkers, somewhat informative was the negative change of thrombocytes (potentially indicative

Table 1 Normalized mutual information (NMI) showing the correlation between individual diagnostic parameters and the composite multimorbidity variable, composed of the presence of ischemic heart disease, COPD, and dementia

No.	Parameter	NMI
1	Cfimdis	0.0954
2	Tfimdis	0.07266
3	Cfimad	0.07138
4	Mfimdis	0.06688
5	Tfimad	0.05406
6	Glucad	0.05372
7	Glucdis	0.0441
8	Mfimad	0.02936
9	Thr − 10%	0.02228
10	Gender	0.02215
11	Gluc − 10%	0.02169
12	Thr + 10%	0.02066
13	Age	0.01914
14	WBCad	0.01842
15	Gluc + 10%	0.01841
16	Natad	0.01803
17	Pulse + 5%	0.01745
18	Pulsedis	0.01569
19	WBC + 1%	0.01458
20	Natdis	0.01388
21	Mfim + 30%	0.01309
22	Cfim + 5%	0.01171
23	Potdis	0.01096
24	Nat − 1%	0.01075
25	Lymad	0.00982
26	Tfim + 30%	0.0097
27	Potad	0.00858
28	Hb > + 15%	0.00803
29	Thrad	0.00803
30	Potad − 5%	0.00765
31	Nat + 1%	0.00748
32	Lym + 10%	0.00735
33	WBC − 1%	0.00715
34	Lymdis	0.00634
35	Pulsead	0.00529
36	Potad + 5%	0.00499
37	Hbad	0.00492
38	Cfim − 5%	0.00489
39	WBCdis	0.00446
40	Thrdis	0.00379
41	Lym − 10%	0.00303
42	Pulse − 5%	0.00157

Table 2 Normalized mutual information (NMI) showing the correlation between combined double diagnostic parameters and the composite multimorbidity variable, composed of the presence of ischemic heart disease, COPD, and dementia (the highest and lowest values are shown)

No.	Parameter 1	Parameter 2	NMI
1	Cfimdis	Glucad	0.15934
2	Cfimdis	Glucdis	0.14889
3	Cfimdis	Mfimdis	0.13497
4	Mfimdis	Glucad	0.1347
5	Cfimdis	Thr − 10%	0.13432
6	Cfimdis	Cfimad	0.13413
7	Tfimdis	Glucad	0.1334
8	Cfimdis	Gender	0.13198
9	Cfimdis	Tfimdis	0.1311
10	Cfimad	Glucad	0.12928
11	Cfimdis	Tfimad	0.12811
12	Cfimdis	Thr + 10%	0.12762
13	Tfimdis	Glucdis	0.12724
14	Cfimdis	Gluc − 10%	0.12516
15	Mfimdis	Glucdis	0.12444
16	Cfimdis	Age	0.12255
17	Cfimad	Glucdis	0.1208
18	Cfimdis	Mfimad	0.12078
19	Tfimad	Glucad	0.11383
20	Tfimdis	Cfimad	0.11343
21	Cfimad	Gender	0.11132
22	Tfimdis	Thr − 10%	0.11095
23	Tfimdis	Tfimad	0.11025
24	Tfimad	Glucdis	0.11023
25	Tfimdis	Gluc − 10%	0.10792
26	Mfimdis	Gluc − 10%	0.1058
27	Tfimdis	Thr + 10%	0.10502
28	Cfimad	Mfimdis	0.10453
29	Tfimdis	Mfimad	0.10333
30	Cfimad	Thr + 10%	0.10227
…	….	…	…
70	Gender	Thr + 10%	0.05875
71	Gender	Age	0.0558
72	Gluc − 10%	Age	0.052
73	Gluc − 10%	Thr + 10%	0.05174
74	Gender	Gluc − 10%	0.05152
75	Thr + 10%	Age	0.05009
76	Thr − 10%	Age	0.04952
77	Thr − 10%	Gluc − 10%	0.04747
78	Thr − 10%	Thr + 10%	0.04335

of the state of the blood clotting system), with the NMI = 0.02228. Gender also provided additional informative value (NMI = 0.02215).

When combining two diagnostic markers, the informative value regarding multimorbidity increased. The NMI correlations between combined double markers and the multimorbidity variable are shown in Table 2. Interestingly, the most informative combinations correlating with multimorbidity combined both functional parameters and a specific biomarker (glucose). Thus, the most informative combined parameter was Cfim at discharge together with glucose at admission (NMI = 0.15934) and Cfim at discharge with glucose at discharge (NMI = 0.14889). Notice the cumulative (synergistic) effect that is produced by such a combination. Thus, for Cfimdis NMI = 0.0954 and for glucose ad, NMI = 0.05372, giving the simple arithmetic sum of NMI = 0.14912, which is less than the cumulative value of the combination of these two parameters: NMI = 0.15934. This may indicate the importance of combining functional measurements with biomarkers measurements.

The combination of different types of functional measurements also increases the informative value (NMI = 0.13497 for Cfmindis together with Mfimdis). The cognitive functional performance appears in the most informative pairs at discharge, after the experience of hip fracture and often painful treatments, possibly indicating the degree of cognitive resilience. Generally, the geriatric cohort under consideration included multiple treatment factors that may complicate the interpretation. Interestingly, age appeared to be of little informative value, either alone or in combinations. This may be characteristic of the particular geriatric sample under consideration, with subjects aged 63–97. Nonetheless, despite those complications, the present methodology permitted to establish good informative values for the correlation of diagnostic parameters and the multimorbidity variable.

Conclusion

The present study offers a methodology to formally describe the multimorbidity variable composed from several age-related diseases, and to find the most informative diagnostic parameters and parameters' combinations for the precise quantitative evaluation of the multimorbidity variable. So far such a formal agreed methodology had been absent, but it is strongly needed

to evaluate elderly geriatric patients, who are as a rule characterized by multimorbidity. It may be particularly helpful to evaluate the effects of aging-ameliorating treatments on aging-derived multimorbidity. The study demonstrated the principal applicability of this methodology in a situation most common for actual clinical geriatric settings. Further analysis of additional diverse clinical data, including data on therapeutic interventions, will improve the clinical utility of such a methodology.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Blokh D, Stambler I (2015a) Information theoretical analysis of aging as a risk factor for heart disease. Aging Dis 6(3):196–207

Blokh D, Stambler I (2015b) Applying information theory analysis for the solution of biomedical data processing problems. Am J Bioinform 3(1):17–29

Blokh D, Stambler I (2016) The application of information theory for the research of aging and aging-related diseases. Prog Neurobiol. https://doi.org/10.1016/j.pneurobio.2016.03.005

Blokh D, Stambler I (2017) The use of information theory for the evaluation of biomarkers of aging and physiological age. Mech Ageing Dev 163:23–29

Dodds TA, Martin DP, Stolov WC, Deyo RA (1993) A validation of the functional independence measurement and its performance among rehabilitation inpatients. Arch Phys Med Rehabil 74:531–536

Fontana L, Kennedy BK, Longo VD, Seals D, Melov S (2014) Medical research: treat ageing. Nature 511(7510):405–407

Goldman DP, Cutler D, Rowe JW, Michaud PC, Sullivan J, Peneva D, Olshansky SJ (2013) Substantial health and economic returns from delayed aging may warrant a new focus for medical research. Health Aff 32(10):1698–1705

Hall SS (2015) A trial for the ages. Science 349(6254):1275–1278

Jin K, Simpkins JW, Ji X, Leis M, Stambler I (2015) The critical need to promote research of aging and aging-related diseases to improve health and longevity of the elderly population. Aging Dis 6(1):1–5

Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB (1994) The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil 75:127–132

Newman JC, Milman S, Hashmi SK, Austad SN, Kirkland JL, Halter JB, Barzilai N (2016) Strategies and challenges in clinical trials targeting human aging. J Gerontol A Biol Sci Med Sci 71(11):1424–1434

Salive ME (2013) Multimorbidity in older adults. Epidemiol Rev 35(1):75–83

Rae MJ, Butler RN, Campisi J, de Grey AD, Finch CE, Gough M, Martin GM, Vijg J, Perrott KM, Logan BJ (2010) The demographic and biomedical case for late-life interventions in aging. Sci Transl Med 2(40):40cm21

Semba RD, Nicklett EJ, Ferrucci L (2010) Does accumulation of advanced glycation end products contribute to the aging phenotype? J Gerontol A Biol Sci Med Sci 65A(9):963–975